

## ON A PROBLEM BY STEKLOV

A. APTEKAREV, S. DENISOV, D. TULYAKOV

ABSTRACT. Given any  $\delta \in (0, 1)$ , we define the Steklov class  $S_\delta$  to be the set of probability measures  $\sigma$  on the unit circle  $\mathbb{T}$ , such that  $\sigma'(\theta) \geq \delta/(2\pi) > 0$  at every Lebesgue point of  $\sigma$ . One can define the orthonormal polynomials  $\phi_n(z)$  with respect to  $\sigma \in S_\delta$ . In this paper, we obtain the sharp estimates on the uniform norms  $\|\phi_n\|_{L^\infty(\mathbb{T})}$  as  $n \rightarrow \infty$  which settles a question asked by Steklov in 1921. As an important intermediate step, we consider the following variational problem. Fix  $n \in \mathbb{N}$  and define  $M_{n,\delta} = \sup_{\sigma \in S_\delta} \|\phi_n\|_{L^\infty(\mathbb{T})}$ . Then, we prove

$$C(\delta)\sqrt{n} < M_{n,\delta} \leq \sqrt{\frac{n+1}{\delta}}.$$

A new method is developed that can be used to study other important variational problems. For instance, we prove the sharp estimates for the polynomial entropy in the Steklov class.

**Introduction.**

One version of the Steklov's problem (see [26], [27]) is to obtain the bounds on the sequence of polynomials  $\{P_n(x)\}_{n=0}^\infty$ , which are orthonormal

$$\int_{-1}^1 P_n(x) P_m(x) \rho(x) dx = \delta_{n,m}, \quad n, m = 0, 1, 2, \dots \quad (0.1)$$

with respect to the strictly positive weight  $\rho$ :

$$\rho(x) \geq \delta > 0, \quad x \in [-1, 1]. \quad (0.2)$$

In 1921, V.A. Steklov made a conjecture that a sequence  $\{P_n(x)\}$  is bounded at any point  $x \in (-1, 1)$ , i.e.,

$$\limsup_{n \rightarrow \infty} |P_n(x)| < \infty \quad (0.3)$$

provided that the weight  $\rho$  does not vanish on  $[-1, 1]$ . On page 321, he writes (adapted translation from French):

*"I believe that inequality (0.3) is the common property of all polynomials whose orthogonality weight  $\rho$  does not vanish inside the given interval, but so far I haven't succeeded in finding either the rigorous proof to that statement or an example when this estimate does not hold at each interior point of the given interval".*

This problem and some related questions gave rise to extensive research, see, e.g., [11, 12, 13, 14] and the survey [27] for a detailed discussion and the list of references. In 1979, Rakhmanov [22] disproved this conjecture by constructing a weight from the Steklov class (0.2), for which

$$\limsup_{n \rightarrow \infty} |P_n(0)| = \infty.$$

It is known (see, for example [10]) that the following bound

$$|P_n(x)| = o(\sqrt{n})$$

holds for any  $x \in (-1, 1)$  as long as  $\rho$  satisfies (0.2). In his next paper [23], Rakhmanov proved that for every  $\varepsilon > 0$  and  $x_0 \in (-1, 1)$  there is a weight  $\rho(x; x_0, \varepsilon)$  from the Steklov class such that the corresponding  $\{P_n(x)\}$  grow as

$$|P_{k_n}(x_0)| \geq k_n^{1/2-\varepsilon}, \quad (0.4)$$

where  $\{k_n\}$  is some subsequence in  $\mathbb{N}$ . In [1], the size of the polynomials was studied for the continuous weight.

All Rakhmanov's counterexamples were obtained as corollaries of the corresponding results for the polynomials  $\{\phi_n\}$  orthonormal on the unit circle

$$\int_0^{2\pi} \phi_n(e^{i\theta}, \sigma) \overline{\phi_m(e^{i\theta}, \sigma)} d\sigma(\theta) = \delta_{n,m}, \quad n, m = 0, 1, 2, \dots, \quad (0.5)$$

$$\phi_n(z, \sigma) = \lambda_n z^n + \dots, \quad \lambda_n > 0$$

with respect to measures from the Steklov class  $S_\delta$  defined as the class of probability measures  $\sigma$  on the unit circle satisfying

$$\sigma' \geq \delta/(2\pi) \quad (0.6)$$

at every Lebesgue point. The version of Steklov's conjecture for this situation would be to prove that the sequence  $\{\phi_n(z, \sigma)\}$  is bounded in  $n$  at every  $z \in \mathbb{T}$  provided that  $\sigma \in S_\delta$ .

This conjecture might be motivated by the following estimate. Consider the Christoffel-Darboux kernel

$$K_n(\xi, z, \mu) = \sum_{j=0}^n \overline{\phi_j(\xi, \mu)} \phi_j(z, \mu)$$

for  $\xi = z$  as the function of  $\mu$ . If  $\mu_1 \leq \mu_2$ , then (see [10] or [24])

$$K_n(z, z, \mu_2) \leq K_n(z, z, \mu_1), \quad z \in \mathbb{T}. \quad (0.7)$$

Here we do not assume  $\mu_{1(2)}$  to be probability measures, of course. Therefore, if  $\sigma \in S_\delta$ , we get

$$K_n(z, z, \sigma) \leq \frac{n+1}{\delta}, \quad \text{i.e.} \quad \frac{1}{n+1} \sum_{j=0}^n |\phi_j(z, \sigma)|^2 \leq \delta^{-1}. \quad (0.8)$$

by taking  $\mu_1 = \delta(2\pi)^{-1} d\theta$  in (0.7).

So, on average the polynomials  $\phi_n(z, \sigma)$  are indeed bounded in  $n$  and one might want to know whether they are bounded for all  $n$ . Rakhmanov proved the following Theorem which gave a negative answer to this question.

**Theorem 0.1.** [23] *Let  $\sigma \in S_\delta$ , where  $\delta$  is sufficiently small. Then, for every sequence  $\{\beta_n\} : \beta_n \rightarrow 0$ , there is  $\sigma \in S_\delta$  such that*

$$\|\phi_{k_n}(z, \sigma)\|_{L^\infty(\mathbb{T})} > \beta_{k_n} \sqrt{\frac{k_n}{\ln^3 k_n}} \quad (0.9)$$

for some sequence  $\{k_n\} \subset \mathbb{N}$ .

This estimate is almost sharp due to the following result (see, e.g., [19], p.11 for the real line case; [10], p. 32, theorem 3.5 for the pointwise estimate).

**Theorem 0.2.** *For  $\sigma \in S_\delta$ , we have*

$$\|\phi_n(z, \sigma)\|_{L^\infty(\mathbb{T})} = o(\sqrt{n}). \quad (0.10)$$

(for completeness, we give the proof in the end of Appendix A).

In the proof of the Theorem 0.1, an important role was played by the following extremal problem. For a fixed  $n$ , define

$$M_{n,\delta} = \sup_{\sigma \in S_\delta} \|\phi_n(z, \sigma)\|_{L^\infty(\mathbb{T})}. \quad (0.11)$$

One of the key results in [23] is the following inequality

$$C \sqrt{\frac{n+1}{\delta \ln^3 n}} \leq M_{n,\delta}, \quad C > 0. \quad (0.12)$$

We recall here a well-known estimate (see [10]):

**Lemma 0.1.** *We have*

$$M_{n,\delta} \leq \sqrt{\frac{n+1}{\delta}}, \quad n \in \mathbb{N}. \quad (0.13)$$

*Proof.* Indeed, this is immediate from the estimate (0.8). We can also argue differently:

$$1 = \int_{\mathbb{T}} |\phi_n|^2 d\sigma \geq \delta/(2\pi) \int_{\mathbb{T}} |\phi_n|^2 d\theta$$

so (0.13) follows from

$$\|\phi_n\|_{L^2(\mathbb{T})}^2 = \left\| \sum_{j=0}^n c_j z^j \right\|_{L^2(\mathbb{T})}^2 = 2\pi \sum_{j=0}^n |c_j|^2 \leq \frac{2\pi}{\delta}$$

and Cauchy-Schwarz

$$\|\phi_n\|_{L^\infty(\mathbb{T})} \leq (n+1)^{1/2} \sqrt{\sum_{j=0}^n |c_j|^2}.$$

□

**Remark.** Notice that all we used in the proof is the normalization  $\|\phi_n\|_{L^2(\mathbb{T}, \sigma)} = 1$  and the Steklov's condition on the measure. The problem, though, is whether the orthogonality leads to further restrictions on the size.

The purpose of the current paper is to obtain the sharp bounds for the problem of Steklov, i.e., the problem of estimating the growth of  $\phi_n$ . We will get rid of the logarithmic factor in the denominator in (0.9) and (0.12) and thus prove the optimal inequalities. The main results are contained in the following two statements:

**Theorem 0.3.** *If  $\delta \in (0, 1)$ , then*

$$M_{n,\delta} > C(\delta) \sqrt{n}. \quad (0.14)$$

and

**Theorem 0.4.** *Let  $\delta \in (0, 1)$ . Then, for every positive sequence  $\{\beta_n\} : \lim_{n \rightarrow \infty} \beta_n = 0$ , there is a probability measure  $\sigma^* : d\sigma^* = \sigma^{*'} d\theta$ ,  $\sigma^* \in S_\delta$  such that*

$$\|\phi_{k_n}(z, \sigma^*)\|_{L^\infty(\mathbb{T})} \geq \beta_{k_n} \sqrt{k_n} \quad (0.15)$$

for some sequence  $\{k_n\} \subset \mathbb{N}$ .

**Remark.** It will be clear later that both results hold for far more regular weights (see Lemma 4.1 and the proof of Theorem 0.4 below).

The Steklov condition (0.6) is quite natural for the analysis of the “size” of the polynomial. Indeed, if  $\mu$  is a measure (not necessarily a probability one) then the trivial scaling

$$\phi_n(z, m\mu) = \frac{\phi_n(z, \mu)}{\sqrt{m}}$$

holds and this changes the size of  $\phi_n$  accordingly. Let now  $\mu$  be a probability measure and  $\Gamma$  be a small arc which does not support all of  $\mu$ . Then, we can take  $\mu_m = \chi_{\Gamma^c} \cdot \mu + m\chi_{\Gamma} \cdot \mu$  with  $m$  very small. So,  $\|\mu_m\| \sim \|\chi_{\Gamma^c} \cdot \mu\| \sim 1$  and one can expect that  $\phi_n(z, m\mu)$  gets large on most of  $\Gamma$  as  $m \rightarrow 0$  in analogy to the case of the whole  $\mathbb{T}$ . This is indeed true for many measures  $\mu$ . Thus, if one studies the dependence of  $\|\phi_n(z, \mu)\|_{L^\infty(\mathbb{T})}$  on  $\mu$ , then the conditions on the measure which control the size of the polynomial should account for that fact and an obvious bound that takes care of this is (0.6) as it does not allow the measure to be scaled on any arc.

The problem of estimating the size of  $\phi_n$  is one of the most basic and most well-studied problems in approximation theory. Nevertheless, the sharp bounds were missing even for the Steklov’s class – the most natural class of measures for this problem. In the current paper, we not only establish these bounds but also suggest a new method which, we believe, is general enough to be used in the study of other variational problems where the constructive information on the weight is given. For example, one can replace the Steklov’s condition by the lower bounds like

$$\sigma'(\theta) \geq w(\theta), \quad \text{for a.e. } \theta \in \mathbb{T}$$

where  $w(\theta)$  vanishes at a point in a particular way (e.g.,  $w(\theta) \sim |\theta|^\alpha$ ). The results we obtain are sharp and we apply them to estimate polynomial entropies – another important quantity to measure the size of the polynomial.

**Remark.** The size, asymptotics, and universality of the Christoffel-Darboux kernel were extensively studied, see, e.g. [17, 16, 18, 28, 29]. We, however, will focus on  $\phi_n$  itself.

**Remark.** Since  $S_\delta$  is invariant under the rotation and  $\{e^{ij\theta_0}\phi_j(ze^{-i\theta_0}, \mu)\}$  are orthonormal with respect to  $\mu(\theta - \theta_0)$ , we can always assume that  $\|\phi_n\|_\infty$  is reached at point  $z = 1$ . Therefore, we have

$$M_{n,\delta} = \sup_{\mu \in S_\delta} |\phi_n(1, \mu)|.$$

One can consider the monic orthogonal polynomials  $\Phi_n(z, \mu) = z^n + \dots$  and the Schur parameters  $\{\gamma_n\}$  so that

$$\phi_n(z, \mu) = \frac{\Phi_n(z, \mu)}{\|\Phi_n\|_{L^2(\mathbb{T}, \mu)}}$$

and

$$\gamma_n = -\overline{\Phi_{n+1}(0, \mu)} \quad (0.16)$$

If  $\rho_n = \sqrt{1 - |\gamma_n|^2}$ , then (see [25])

$$\Phi_n(z, \mu) = \phi_n(z, \mu)\lambda_n^{-1}, \quad \lambda_n = \left(\rho_0 \cdot \dots \cdot \rho_{n-1}\right)^{-1}. \quad (0.17)$$

The Szegő formula [25] yields

$$\exp\left(\frac{1}{4\pi} \int_{-\pi}^{\pi} \ln(2\pi\mu'(\theta))d\theta\right) = \prod_{j \geq 0} \rho_j. \quad (0.18)$$

So, for  $\mu \in S_\delta$ , we have

$$\sqrt{\delta} \leq \prod_{n \geq 0} \rho_n \leq 1$$

and therefore

$$\sqrt{\delta} |\phi_n(z, \mu)| \leq |\Phi_n(z, \mu)| \leq |\phi_n(z, \mu)|, \quad z \in \mathbb{C}$$

for any  $\mu \in S_\delta$ . Thus, we have

$$\delta^{1/2} M_{n,\delta} \leq \sup_{\mu \in S_\delta} \|\Phi_n(z, \mu)\|_{L^\infty(\mathbb{T})} \leq M_{n,\delta} \quad (0.19)$$

and for fixed  $\delta$  the variational problems for orthonormal and monic orthogonal polynomials are equivalent.

The estimate (0.13) can not possibly be sharp for  $\delta$  very close to 1. Indeed, if  $\delta = 1$  then  $\sigma$  is the Lebesgue measure and  $\phi_n(z) = z^n$ . We have the following result which provides an effective bound and improves (0.13) for  $\delta$  close to 1.

**Lemma 0.2.** *We have*

$$M_{n,\delta} \leq \delta^{-1/2} \left( 1 + \sqrt{\frac{n(1-\delta)}{\delta}} \right)$$

*Proof.* Let  $\tilde{\sigma}$  be one of the maximizers for  $M_{n,\delta}$ , i.e.,  $\|\phi_n(z, \tilde{\sigma})\|_{L^\infty(\mathbb{T})} = M_{n,\delta}$ . The existence of such a maximizer is proved in Theorem 1.1 below. Then,  $d\tilde{\sigma} = (2\pi)^{-1} \delta d\theta + d\tilde{\mu}$  where  $\|\tilde{\mu}\| = 1 - \delta$ . Let  $\Phi_n(z, \tilde{\sigma})$  be the corresponding monic polynomial. We use the variational characterization of  $\Phi_n$  (see [24]):  $\Phi_n = \arg \min_{P(z)=z^n+\dots} \|P\|_{L^2(\mathbb{T}, \tilde{\sigma})}^2$ . If  $\Phi_n = z^n + \tilde{a}_{n-1}z^{n-1} + \dots + \tilde{a}_1z + \tilde{a}_0$ , then

$$\tilde{a} = \arg D_{n,\delta}, \quad D_{n,\delta} = \min_{a=(a_0, \dots, a_{n-1})} \left( \delta \|a\|_{\ell^2}^2 + \int_{\mathbb{T}} |z^n + a_{n-1}z^{n-1} + \dots + a_1z + a_0|^2 d\tilde{\mu} \right)$$

In particular, upon choosing  $a = 0$ , we get  $D_{n,\delta} \leq \int_{\mathbb{T}} d\tilde{\mu} = 1 - \delta$  and so  $\|\tilde{a}\|_{\ell^2}^2 \leq (1 - \delta)/\delta$ . Then, Cauchy-Schwarz inequality gives

$$\|\Phi_n\|_{L^\infty(\mathbb{T})} \leq 1 + \sqrt{\frac{n(1-\delta)}{\delta}}$$

and (0.19) finishes the proof.  $\square$

Now, we would like to comment a little on the methods we use. The proofs by Rakhmanov were based on the following formula for the orthogonal polynomial that one gets after adding several point masses to a “background” measure at particular locations on the circle (see [22]).

**Lemma 0.3.** *Let  $\mu$  be a positive measure on  $\mathbb{T}$ ,  $\Phi_n(z, \mu)$  be the corresponding monic orthogonal polynomials, and*

$$K_n(\xi, z, \mu) = \sum_{l=0}^n \overline{\phi_j(\xi, \mu)} \phi_j(z, \mu)$$

*be the Christoffel-Darboux kernel, i.e.*

$$P(\xi) = \langle P(z), K_n(\xi, z, \mu) \rangle_{L^2(\mathbb{T}, \mu)}, \quad \deg P \leq n.$$

*Then, if  $\xi_j \in \mathbb{T}, j = 1, \dots, m, m \leq n$  are chosen such that*

$$K_{n-1}(\xi_j, \xi_l, \mu) = 0, \quad j \neq l \quad (0.20)$$

*then*

$$\Phi_n(z, \eta) = \Phi_n(z, \mu) - \sum_{k=1}^m \frac{m_k \Phi_n(\xi_k, \mu)}{1 + m_k K_{n-1}(\xi_k, \xi_k, \mu)} K_{n-1}(\xi_k, z, \mu) \quad (0.21)$$

where

$$\eta = \mu + \sum_{k=1}^m m_k \delta_{\theta_k}, \quad z_k = e^{i\theta_k}, \quad m_k \geq 0.$$

The limitation that  $\xi_j$  must be the roots of  $K$  is quite restrictive and the direct application of this formula with background  $d\mu = d\theta$  yields logarithmic growth at best. In the later paper [23], Rakhmanov again ingeniously used the idea of inserting the point mass but the resulting bound (0.12) contained the logarithm in the denominator and the measure of orthogonality was not defined explicitly.

We will use a completely different approach. First, we will rewrite the Steklov condition in the convenient form as some estimate that involves Caratheodory function and a polynomial (see Lemma 3.3 below). This decoupling is basically equivalent to solving the well-known truncated trigonometric moments problem. Then, we will present a particular function and a polynomial and show that they satisfy the necessary conditions. This allows us to have a good control on the size of the polynomial itself and on the structure of the measure of orthogonality.

The paper has four parts and two Appendixes. The first part contains results on the structure of an optimal measure and discussion of the case when  $\delta$  is  $n$ -dependent and very small. In the second part, the proof of Theorem 0.3 is given for fixed small  $\delta$ . We will apply the “localization principle” to handle every  $\delta \in (0, 1)$  and prove Theorem 0.4 in the third part. In the last one, two applications are given. First, the lower bounds are obtained for polynomials orthogonal on the real line. Then, we prove the sharp estimates for the polynomial entropies in the Steklov class. The Appendixes contain some auxiliary results we use in the main text.

Here are some notation used in the paper: the Cauchy kernel for the unit circle is denoted by  $C(z, \xi)$ , i.e.

$$C(z, \xi) = \frac{\xi + z}{\xi - z}, \quad \xi \in \mathbb{T}.$$

If the function is analytic in  $\mathbb{D}$  and has a nonnegative real part there, then we will call it Caratheodory function.

Given any polynomial  $P_n(z) = p_n z^n + \dots + p_1 z + p_0$ , we can define its  $n$ -th reciprocal (or the  $*$ -transform)

$$P_n^*(z) = z^n \overline{P_n(1/\bar{z})} = \bar{p}_0 z^n + \bar{p}_1 z^{n-1} + \dots + \bar{p}_n.$$

Notice that if  $z^*$  is a root of  $P_n(z)$  and  $z^* \neq 0$ , then  $(\bar{z}^*)^{-1}$  is a root of  $P_n^*(z)$ .

Given two positive functions  $F_1$  and  $F_2$  defined on  $\mathbb{D}$ , we write  $F_1 \lesssim F_2$  if there is a constant  $C$  (that might depend only on the fixed parameters) such that

$$F_1 < C F_2, \quad C > 0$$

on  $\mathbb{D}$ . We write  $F_1 \sim F_2$  if

$$F_1 \lesssim F_2 \lesssim F_1.$$

We use the notation  $b = O^*(a)$  if  $b \sim a$ . The symbol  $\delta_a$  denotes the delta function (the point mass) supported at  $a \in (-\pi, \pi]$  or at complex point  $e^{ia} \in \mathbb{T}$ . If  $\varepsilon$  is a positive parameter, then  $\varepsilon \ll 1$  is the shorthand for: “ $\varepsilon < \varepsilon_0$ , where  $\varepsilon_0$  is sufficiently small”. If  $p(z) = a_n z^n + \dots + a_1 z + a_0$ , then we define  $\text{coeff}(p, j) = a_j$ .

We will use the following standard notation for the norms. If  $\mu$  is a measure,  $\|\mu\|$  refers to its total variation. For functions  $f$  defined on  $[-\pi, \pi]$ , we write

$$\|f\|_p = \|f\|_{L^p[-\pi, \pi]}, \quad 1 \leq p \leq \infty.$$

The symbol  $\langle f, g \rangle_\sigma$  denotes the following inner product given by

$$\langle f, g \rangle_\sigma = \int_M f(x) \overline{g}(x) d\sigma$$

where  $\sigma$  is a measure on  $M$  (e.g.,  $M = \mathbb{T}$  or  $M = [-\pi, \pi]$ ).

## Part 1. Variational problem: structure of the extremizers

### 1. STRUCTURE OF THE EXTREMAL MEASURE.

In this section, we first address the problem of the existence of maximizers, i.e.,  $\mu_n^* \in S_\delta$  for which

$$M_{n,\delta} = |\phi_n(1; \mu_n^*)|. \quad (1.1)$$

We will prove that these extremizers exist and will study their properties.

**Theorem 1.1.** *There are  $\mu_n^* \in S_\delta$  for which (1.1) holds.*

*Proof.* Suppose  $\mu_k \in S_\delta$  is the sequence which yields the sup, i.e.

$$|\phi_n(1, \mu_k)| \rightarrow M_{n,\delta}, \quad k \rightarrow \infty.$$

Since the unit ball is weak-(\*) compact, we can choose  $\mu_{k_j} \rightarrow \mu^*$  and this convergence is weak-(\*), i.e.

$$\int f d\mu_{k_j} \rightarrow \int f d\mu^*, \quad j \rightarrow \infty$$

for any  $f \in C(\mathbb{T})$ . In particular,  $\mu^*$  is a probability measure. Moreover, for any interval  $(a, b) \subseteq (-\pi, \pi]$ , we have (assuming, e.g., that the endpoints  $a$  and  $b$  are not atoms for  $\mu^*$ ):

$$\int_{[a,b]} d\mu^* \geq \delta(b-a)/(2\pi)$$

since each  $\mu_{k_j} \in S_\delta$ . This implies  $\mu^{*'} \geq \delta/(2\pi)$  a.e. on  $\mathbb{T}$ . The moments of  $\mu_{k_j}$  will converge to the moments of  $\mu^*$  and therefore

$$\phi_n(1, \mu_{k_j}) \rightarrow \phi_n(1, \mu^*).$$

Therefore,  $\mu^* \in S_\delta$  and  $|\phi_n(1, \mu^*)| = M_{n,\delta}$ . □

This argument gives existence of an extremizer. Although we do not know whether it is unique, we can prove that every  $d\mu^*$  must have a very special form.

**Theorem 1.2.** *If  $\mu^*$  is a maximizer then it can be written in the following form*

$$d\mu^* = (2\pi)^{-1} \delta d\theta + \sum_{j=1}^N m_j \delta_{\theta_j}, \quad 1 \leq N \leq n \quad (1.2)$$

where  $m_j \geq 0$  and  $-\pi < \theta_1 < \dots < \theta_N \leq \pi$ .

Suppose we have a positive measure  $\mu$  and its moments are given by

$$s_j = \int e^{ij\theta} d\mu = s_j^R + i s_j^I, \quad j = 0, 1, \dots$$

Then, the following formulas are well-known ([25])

$$\Phi_n(z, \sigma) = \mathfrak{D}_{n-1}^{-1} \begin{vmatrix} s_0 & s_1 & \dots & s_n \\ \bar{s}_1 & s_0 & \dots & s_{n-1} \\ \dots & \dots & \dots & \dots \\ \bar{s}_{n-1} & \bar{s}_{n-2} & \dots & s_1 \\ 1 & z & \dots & z^n \end{vmatrix} \quad (1.3)$$

$$\mathfrak{D}_n = \det \mathfrak{T}_n, \quad \mathfrak{T}_n = \begin{bmatrix} s_0 & s_1 & \dots & s_n \\ \bar{s}_1 & s_0 & \dots & s_{n-1} \\ \dots & \dots & \dots & \dots \\ \bar{s}_{n-1} & \bar{s}_{n-2} & \dots & s_1 \\ \bar{s}_n & \bar{s}_{n-1} & \dots & s_0 \end{bmatrix}. \quad (1.4)$$

These identities show that  $\Phi_n(z, \sigma)$  depends only on the first  $n$  moments of the measure  $\sigma$ :  $\Phi_n(z, \sigma) = \Phi_n(z, s_0, \dots, s_n)$ . Moreover, by definition of the monic orthogonal polynomial,

$$\Phi_n(z, s_0, s_1, \dots, s_n) = \Phi_n(z, 1, s_1/s_0, \dots, s_n/s_0),$$

i.e.,  $\Phi_n(z, \sigma)$  does not depend on the normalization of the measure.

The functions  $F_{1(2)}$ , given by

$$F_1(s) = |\Phi_n(1, s_0, \dots, s_n)|^2, \quad F_2(s) = |\phi_n(1, s_0, \dots, s_n)|^2, \quad s = (s_0, s_1^R, \dots, s_n^I) \in \mathbb{R}^{2n+1}$$

are the smooth functions of the variables  $\{s_0, s_j^R, s_j^I, j = 1, \dots, n\}$  wherever they are defined. Consider  $\Omega_n = \{s : \mathfrak{T}_n(s) > 0\}$ . If  $s \in \Omega_n$ , then there is a family of measures  $\mu$  which have  $(s_0, \dots, s_n)$  as the first  $n$  moments. That follows from the solution to the truncated trigonometric moment problem.

We will need the following

**Lemma 1.1.** *The functions  $F_{1(2)}(s)$  do not have stationary points on  $\Omega_n$ .*

*Proof.* It is known [25] that the map between the first  $n$  Schur parameters (see (0.16)) and the first  $n$  moments of a probability measure, i.e.,  $\{\gamma_j\}_{j=0}^{n-1} \in \mathbb{D}^n \rightarrow (1, s_1, \dots, s_n)$ , is a bijection. The formulas (0.16) and (1.3) imply that

$$\gamma_{n-1} = \bar{s}_n \frac{\mathfrak{D}_{n-2}}{\mathfrak{D}_{n-1}} + f(s_0, s_1, \bar{s}_1, \dots, s_{n-1}, \bar{s}_{n-1}). \quad (1.5)$$

The both polynomials  $\Phi_n$  and  $\phi_n$  satisfy the recurrences ([25], p.57)

$$\Phi_n(1, \sigma) = \Phi_{n-1}(1, \sigma) - \bar{\gamma}_{n-1} \Phi_{n-1}^*(1, \sigma)$$

and

$$\phi_n(1, \sigma) = \rho_{n-1}^{-1} (\phi_{n-1}(1, \sigma) - \bar{\gamma}_{n-1} \phi_{n-1}^*(1, \sigma)),$$

which shows that

$$\nabla_{s_n} |\Phi_n(1, \sigma)|^2 = \left( \frac{\partial |\Phi_n(1, \sigma)|^2}{\partial s_n^R}, \frac{\partial |\Phi_n(1, \sigma)|^2}{\partial s_n^I} \right) \neq 0,$$

because  $\Phi_{n-1}(z, \sigma)$  and  $\Phi_{n-1}^*(z, \sigma)$  do not depend on  $s_n$ ,  $\Phi_{n-1}^*(1, \sigma) \neq 0$ , and  $\Phi_n(1, \sigma) \neq 0$ . Since  $\rho_{n-1} = \sqrt{1 - |\gamma_{n-1}|^2}$ , we have

$$|\phi_n(1, \sigma)|^2 = C \frac{|\xi - \gamma_{n-1}|^2}{1 - |\gamma_{n-1}|^2}, \quad \xi = \frac{\bar{\phi}_{n-1}(1, \sigma)}{\phi_{n-1}^*(1, \sigma)}, \quad |\xi| = 1, \quad C = |\phi_{n-1}^*(1, \sigma)|^2.$$



We can rewrite it as

$$|\phi_n(1, \sigma)|^2 = C \frac{1 + |\gamma_{n-1}|^2 - 2 \operatorname{Re}(\gamma_{n-1} \bar{\xi})}{1 - |\gamma_{n-1}|^2},$$

which shows that

$$\nabla_{\gamma_{n-1}} |\phi_n(1, \sigma)|^2 \neq 0.$$

where  $|\phi_n(1, \sigma)|^2$  is considered as a function of  $\{\gamma_0, \dots, \gamma_{n-1}\}$ .

Now (1.5) yields

$$\nabla_{s_n} |\phi_n(1, \sigma)|^2 \neq 0$$

and the proof is finished.  $\square$

**Remark.** The proof actually shows that  $\nabla_{s_n} F_{1(2)} \neq 0$ .

*Proof. (of the Theorem 1.2)* Our variational problem is an extremal problem for a functional  $F(s_0, s_1^R, \dots, s_n^I)$  on the finite number of moments  $\{s_0, s_1^R, \dots, s_n^I\}$  of a measure  $\mu$  from  $S_\delta$ . We can take

$$F(s_0, s_1^R, \dots, s_n^I) = |\phi_n(1, s_0, \dots, s_n)|^2.$$

The function  $F$  is differentiable. Moreover,

$$s_0 = \int d\mu, \quad s_j^R = \int \cos(j\theta) d\mu, \quad s_j^I = \int \sin(j\theta) d\mu.$$

Considering the moments as functionals of  $\mu$ , we compute the derivative of  $F$  at the point  $\mu^*$  in the direction  $\delta\mu$ :

$$dF = \int \left( \frac{\partial F}{\partial s_0}(s^*) + \frac{\partial F}{\partial s_1^R}(s^*) \cos(\theta) + \dots + \frac{\partial F}{\partial s_n^I}(s^*) \sin(n\theta) \right) d(\delta\mu).$$

Consider the trigonometric polynomial:

$$T_n(\theta) = \frac{\partial F}{\partial s_0}(s^*) + \frac{\partial F}{\partial s_1^R}(s^*) \cos(\theta) + \dots + \frac{\partial F}{\partial s_n^I}(s^*) \sin(n\theta).$$

From the previous Lemma and Remark, we know that it has degree  $n$ . Let  $M = \max T_n(\theta)$  and  $\{\theta_j; j = 1, \dots, N\}$  are all points where  $M$  is achieved. Clearly,  $N \leq n$ .

Now, if we find a smooth curve  $\mu(t)$ ,  $t \in (0, 1]$  such that  $\mu(t) \in S_\delta$ ,  $\mu(1) = \mu^*$  and define

$$H(t) = F(s_0(\mu(t)), s_1^R(\mu(t)), \dots, s_n^I(\mu(t))),$$

then  $H'(1) \geq 0$  as follows from the optimality of  $\mu^*$ .

Now, we will assume that the measure  $\mu^*$  is not of the form (1.2) and then will come to a contradiction by choosing the curve  $\mu(t)$  in a suitable way.

We will first prove that the singular part of  $\mu^*$  can be supported only at the points  $\{\theta_j\}$ . Indeed, suppose we have

$$\mu^* = \mu_1 + \mu_2$$

where  $\mu_2$  is singular and supported away from  $\{\theta_j\}$ . Consider two smooth functions  $p_1(t)$  and  $p_2(t)$  defined on  $(0, 1]$  that satisfy

$$\|\mu_1\| + p_1(t) + p_2(t)\|\mu_2\| = 1, \quad p_{1(2)}(t) \geq 0, \quad p_1(1) = 0, \quad p_2(1) = 1.$$

For example, one can take  $p_1(t) = \|\mu_2\|(1 - t)$ ,  $p_2(t) = t$ . Take  $\mu(t) = \mu_1 + p_1(t)\delta_{\theta_1} + p_2(t)\mu_2$ . We have  $\mu(t) \in S_\delta$  and

$$H'(1) = \int T_n(\theta) d\mu_2 - \|\mu_2\| T_n(\theta_1) < 0$$

since  $\theta_1$  is a point of global maximum for  $T_n$  and  $\mu_2$  is supported away from  $\{\theta_j\}$  by assumption. This contradicts optimality of  $\mu^*$  and so  $\mu_2 = 0$ .

We can prove similarly now that  $(\mu^*)' = (2\pi)^{-1}\delta$  a.e. Indeed, suppose

$$\mu^* = \mu_1 + \mu_2, \quad d\mu_2 = f(\theta)X_\Omega d\theta$$

where  $f(\theta) > (2\pi)^{-1}\delta_1 > (2\pi)^{-1}\delta$  on  $\Omega$ ,  $|\Omega| > 0$  and  $\mu_1$  is supported on  $\Omega^c$ . We consider the curve

$$\mu(t) = \mu_1 + p_1(t)\delta_{\theta_1} + p_2(t)\mu_2(t) .$$

The choice of  $p_{1(2)}$  is the same. Then,  $\mu(t) \in S_\delta$  for  $t \in (1-\varepsilon, 1)$  provided that  $\varepsilon(\delta_1)$  is small. The similar calculation yields  $H'(1) < 0$  and that gives a contradiction.  $\square$

Since the maximizer in the Steklov problem is given by (1.2), we want to make an observation. The following result is attributed to Geronimus (see [10]).

**Lemma 1.2.** *Consider  $\mu(t) = (1-t)\mu + t\delta_\beta$  where  $t \in (0, 1), \beta \in [-\pi, \pi)$ . Then,*

$$\Phi_n(z, \mu(t)) = \Phi_n(z, \mu) - t \frac{\Phi_n(\xi, \mu)K_{n-1}(\xi, z, \mu)}{1-t+tK_{n-1}(\xi, \xi, \mu)}, \xi = e^{i\beta}. \quad (1.6)$$

*Proof.* Notice that the right hand side is a monic polynomial of degree  $n$ . Then,

$$\langle \text{r.h.s.}, z^j \rangle_{\mu(t)} = 0, \quad j = 0, \dots, n-1$$

which yields orthogonality.  $\square$

The formula (1.6) expresses monic polynomials obtained by adding one point mass to an arbitrary measure at any location. One can try to iterate it to get the optimal measure  $d\mu^*$ . That, however, leads to very complicated analysis.

## 2. THE REGIME OF SMALL $n$ -DEPENDENT $\delta$ .

One can make a trivial observation that if  $\mu$  is any positive measure (not necessarily a probability one) and  $\phi_n(z, \mu)$  is the corresponding orthonormal polynomial, then

$$\phi_n(z, \alpha\mu) = \alpha^{-1/2}\phi_n(z, \mu) \quad (2.1)$$

for every  $\alpha > 0$ . The monic orthogonal polynomials, though, stay unchanged

$$\Phi_n(z, \alpha\mu) = \Phi_n(z, \mu) .$$

Now, consider the modification of the problem: we define

$$\widetilde{M}_{n,\delta} = \sup_{\mu' \geq \delta/(2\pi)} \|\phi_n(z, \mu)\|_{L^\infty(\mathbb{T})} = \sup_{\mu' \geq \delta/(2\pi)} |\phi_n(1, \mu)|,$$

i.e., we drop the requirement for the measure  $\mu$  to be a probability measure. In this case, the upper estimate for  $\widetilde{M}_{n,\delta}$  stays the same and the proof of

$$\widetilde{M}_{n,\delta} \leq \sqrt{\frac{n+1}{\delta}} .$$

is identical. It turns out that the sharp lower bound can be easily obtained in this case.

**Theorem 2.1.** *We have*

$$\widetilde{M}_{n,\delta} = \sqrt{\frac{n+1}{\delta}} .$$

*Proof.* Consider

$$d\sigma = \frac{\delta}{2\pi} d\theta + \sum_{k=1}^n m_k \delta_{\theta_k}, \quad \theta_k = \frac{k}{n+1} 2\pi, \quad k = 1, \dots, n. \quad (2.2)$$

We assume that all  $m_k \geq 0$ . Consider

$$\Pi_n(z) = \prod_{k=1}^n (z - \varepsilon_k), \quad \varepsilon_k = e^{i\theta_k}. \quad (2.3)$$

One gets:  $\Pi_n(z) = 1 + z + \dots + z^n$ ,  $\|\Pi_n\|_{L^2(\mathbb{T}, \sigma)}^2 = \delta(n+1)$ . We define now

$$\Phi_n = \Pi_n + Q_{n-1}$$

where  $Q_{n-1}(z) = q_{n-1}z^{n-1} + \dots + q_1z + q_0$  is chosen to guarantee the orthogonality  $\langle \Phi_n, z^j \rangle_\sigma = 0$ ,  $j = 0, \dots, n-1$ . Suppose that  $m_k = m$  for all  $k$ . Then, we have the following equations

$$\delta + \delta q_j + m \sum_{l=0}^{n-1} q_l \sum_{k=1}^n \varepsilon_k^{l-j} = 0, \quad j = 0, \dots, n-1.$$

Then, since

$$\sum_{k=0}^n \varepsilon_k^d = 0, \quad d \in \{-n, \dots, -1, 1, \dots, n\}$$

we get

$$\delta + \delta q_j + m(n+1)q_j - m \sum_{l=0}^{n-1} q_l = 0, \quad j = 0, \dots, n-1$$

and

$$q_j = -\frac{\delta}{\delta + m}, \quad j = 0, \dots, n-1.$$

Thus,

$$\Phi_n(z) = \Phi_n(z, \sigma) = 1 + \dots + z^n - \frac{\delta}{\delta + m}(1 + \dots + z^{n-1}) = \frac{m}{\delta + m}\Pi_n(z) + \frac{\delta}{\delta + m}z^n.$$

Now, we have

$$\|\Phi_n\|_{L^\infty(\mathbb{T})} = \Phi_n(1, \sigma) = n + 1 - \frac{\delta}{\delta + m}n = 1 + \frac{mn}{\delta + m}$$

and

$$\|\Phi_n\|_{L^2(\mathbb{T}, \sigma)}^2 = \delta \left( 1 + \frac{m^2 n}{(\delta + m)^2} \right) + \frac{\delta^2 n m}{(\delta + m)^2} = \delta \left( 1 + \frac{mn}{\delta + m} \right).$$

For the orthonormal polynomial,

$$\phi_n(1, \sigma) = \frac{\Phi_n(1, \sigma)}{\|\Phi_n\|_{L^2(\mathbb{T}, \sigma)}}.$$

For fixed  $n$ , this gives

$$\lim_{m \rightarrow \infty} \|\phi_n\|_{L^\infty(\mathbb{T})} = \sqrt{(n+1)/\delta}.$$

□

**Remark.** This Theorem has the following implication for our original problem. Suppose we consider the class  $S_\delta$  but  $\delta$  is small in  $n$ . Then, (2.1) gives

$$M_{n,\delta_n} = \sqrt{\frac{n+1}{\delta_n}}(1 + o(1))$$

where

$$\delta_n = \frac{C}{nm_n}, \quad m_n \rightarrow +\infty, \quad \text{as } n \rightarrow \infty.$$

Thus, for  $\delta$  small in  $n$ , the upper bound for  $M_{n,\delta}$  is sharp. If one takes  $m_n = 1/n$  in the proof above to make the total mass finite, the polynomials  $\phi_n$  are bounded in  $n$  as  $\delta \sim 1$ .

## Part 2. The proof of Theorem 0.3: the case of small fixed $\delta$

### 3. LOWER BOUNDS: FIXED $\delta$ AND $n$ .

In this section, we prove the sharp lower bound for small fixed  $\delta$ . The main result is the following Theorem.

**Theorem 3.1.** *There is  $\delta_0 \in (0, 1)$  such that*

$$M_{n,\delta_0} \gtrsim \sqrt{n}. \quad (3.1)$$

**Remark.** In this section, we are not trying to control the size of  $\delta_0$ . The full range  $\delta \in (0, 1)$  will be covered in part 3 by using certain localization technique.

**3.1. Notation and Basics from the theory of polynomials orthogonal on the circle.** We start by introducing some notation and recalling the relevant facts from the theory of polynomials orthogonal on the unit circle.

The following trivial Lemma will be needed later (see, e.g., [21], p. 108)

**Lemma 3.1.** *Let  $n \geq 1$ . If a polynomial  $P_n$  of degree at most  $n$  has all zeroes outside  $\overline{\mathbb{D}}$ , then  $D_n(z) = P_n(z) + P_n^*(z)$  has all (exactly  $n$ ) zeroes on the unit circle.*

*Proof.* We have

$$D_n(z) = P_n \left( 1 + \frac{P_n^*}{P_n} \right), \quad z \in \mathbb{D}.$$

The first factor has no zeroes in  $\mathbb{D}$ . In the second one,  $P_n^*/P_n$  is a Blaschke product (indeed,  $|P_n^*/P_n| = 1$  on  $\mathbb{T}$ ). So,  $1 + \frac{P_n^*}{P_n}$  is holomorphic in  $\mathbb{D}$ , continuous up to the boundary, and its boundary values belong to the circle with center at  $z = 1$  and radius 1. Thus,

$$\operatorname{Re} \left( 1 + \frac{P_n^*}{P_n} \right) \geq 0.$$

Therefore,

$$\text{either} \quad \operatorname{Re} \left( 1 + \frac{P_n^*(z)}{P_n(z)} \right) > 0 \quad (z \in \mathbb{D}) \quad \text{or} \quad 1 + \frac{P_n^*(z)}{P_n(z)} \equiv 0.$$

The last condition implies  $P_n = -P_n^*$ , so  $P_n \neq 0$  on  $\overline{\mathbb{D}}$  by assumption of the Lemma and  $P_n \neq 0$  on  $\overline{\mathbb{C}} \setminus \mathbb{D}$  because it is equal to  $-P_n^*$ . Then,  $P_n \neq 0$  on  $\overline{\mathbb{C}}$  and so  $P_n \equiv \text{const}$ . Therefore  $P_n^* = -P_n \equiv \text{const}$ . This is possible for  $n = 0$  only. Finally,

$$1 + \frac{P_n^*}{P_n}$$

does not have zeroes in  $\mathbb{D}$ . Since  $D_n$  is invariant under the  $*$ -transform, it has the following property:  $D_n(w) = 0$  implies  $D_n(\overline{w}^{-1}) = 0$ . Therefore,  $D_n$  has no zeroes in  $|z| > 1$  as well.

One can actually show that  $D_n$  has the degree  $n$  under the assumptions of the Lemma. Indeed, if  $D_n(z) \neq 0$ ,  $z \in \mathbb{D}$ , then  $\text{Var arg } D_n|_{r\mathbb{T}} = 0$ ,  $r < 1$ . Therefore  $\text{Var arg } D_n^*|_{\frac{1}{r}\mathbb{T}} = 2\pi n \implies \deg(D_n^*) \geq n$ . But  $D_n^* = D_n$ .  $\square$

We will be mostly working with the orthonormal polynomials  $\phi_n$  and the corresponding  $\phi_n^*$ . It is well known [25] that all zeroes of  $\phi_n$  are inside  $\mathbb{D}$  thus  $\phi_n^*$  has no zeroes in  $\overline{\mathbb{D}}$ . However, we also need to introduce the second kind polynomials  $\psi_n$  along with the corresponding  $\psi_n^*$ . Let us recall ([25], p. 57) that

$$\begin{cases} \phi_{n+1} = \rho_n^{-1}(z\phi_n - \overline{\gamma}_n\phi_n^*), & \phi_0 = \sqrt{1/|\mu|} = 1 \\ \phi_{n+1}^* = \rho_n^{-1}(\phi_n^* - \gamma_n z\phi_n), & \phi_0^* = \sqrt{1/|\mu|} = 1 \end{cases} \quad (\text{probability case}) \quad (3.2)$$

and the second kind polynomials satisfy the recursion with Schur parameters  $-\gamma_n$ , i.e.,

$$\begin{cases} \psi_{n+1} = \rho_n^{-1}(z\psi_n + \overline{\gamma}_n\psi_n^*), & \psi_0 = \sqrt{|\mu|} = 1 \\ \psi_{n+1}^* = \rho_n^{-1}(\psi_n^* + \gamma_n z\psi_n), & \psi_0^* = \sqrt{|\mu|} = 1 \end{cases} \quad (3.3)$$

The following Bernstein-Szegő approximation result is valid:

**Theorem 3.2.** *Suppose  $d\mu$  is a probability measure and  $\{\phi_j\}$  and  $\{\psi_j\}$  are the corresponding orthonormal polynomials of the first/second kind, respectively. Then, for any  $N$ , the function*

$$F_N(z) = \frac{\psi_N^*(z)}{\phi_N^*(z)} = \int_{\mathbb{T}} C(z, e^{i\theta}) d\mu_N(\theta), \quad d\mu_N(\theta) = \frac{d\theta}{2\pi|\phi_N(e^{i\theta})|^2} = \frac{d\theta}{2\pi|\phi_N^*(e^{i\theta})|^2}$$

*has the first  $N$  Taylor coefficients identical to the Taylor coefficients of the function*

$$F(z) = \int_{\mathbb{T}} C(z, e^{i\theta}) d\mu(\theta).$$

*In particular, the polynomials  $\{\phi_j\}$  and  $\{\psi_j\}$ ,  $j \leq N$  are the orthonormal polynomials of the first/second kind for the measure  $d\mu_N$ .*

We also need the following Lemma:

**Lemma 3.2.** *The polynomial  $P_n(z)$  of degree  $n$  is the orthonormal polynomial for a probability measure with infinitely many growth points if and only if*

1.  $P_n(z)$  has all  $n$  zeroes inside  $\mathbb{D}$  (counting the multiplicities).
2. The normalization conditions

$$\int_{\mathbb{T}} \frac{d\theta}{2\pi|P_n(e^{i\theta})|^2} = 1, \quad \text{coeff}(P_n, n) > 0$$

*are satisfied.*

Now, we are ready to formulate the main result of this section.

**3.2. The reduction of the problem: Decoupling Lemma.** The proof of the Theorem 3.1 will be based on the following result.

**Lemma 3.3. (The Decoupling Lemma)** *To prove (3.1), it is sufficient to find a polynomial  $\phi_n^*$  and a Caratheodory function  $\tilde{F}$  which satisfy the following properties:*

1.  $\phi_n^*(z)$  has no roots in  $\mathbb{D}$ .
2. Normalization on the size and “rotation”

$$\int_{\mathbb{T}} |\phi_n^*(z)|^{-2} d\theta = 2\pi, \quad \phi_n^*(0) > 0. \quad (3.4)$$

3. Large uniform norm, i.e.,

$$|\phi_n^*(1)| \sim \sqrt{n}.$$

4.  $\tilde{F} \in C^\infty(\mathbb{T})$ ,  $\operatorname{Re} \tilde{F} > 0$  on  $\mathbb{T}$ , and

$$\frac{1}{2\pi} \int_{\mathbb{T}} \operatorname{Re} \tilde{F}(e^{i\theta}) d\theta = 1. \quad (3.5)$$

5. Moreover,

$$|\phi_n^*(z)| + |\tilde{F}(z)(\phi_n(z) - \phi_n^*(z))| < C_1(\delta) \left( \operatorname{Re} \tilde{F}(z) \right)^{1/2} \quad (3.6)$$

uniformly in  $z \in \mathbb{T}$ .

*Proof.* By Lemma 3.2, the first two conditions guarantee that  $\phi_n(z)$  is an orthonormal polynomial of some probability measure. It also determines the first  $n$  Schur parameters:  $\gamma_0, \dots, \gamma_{n-1}$ . The third one gives the necessary growth. Next, let us show that the fourth and fifth conditions are sufficient for the existence of a measure  $\sigma \in S_\delta$  for which  $\phi_n$  is the  $n$ -th orthonormal polynomial.

By the fourth condition,  $\tilde{F}$  defines the probability measure  $\tilde{\sigma}$  which is purely absolutely continuous and has positive smooth density  $\tilde{\sigma}'$  given by

$$\tilde{\sigma}'(\theta) = \frac{\operatorname{Re} \tilde{F}(e^{i\theta})}{2\pi}. \quad (3.7)$$

Denote its Schur parameters by  $\{\tilde{\gamma}_j\}$ ,  $j = 0, 1, \dots$  and the orthonormal polynomials of the first and second kind by  $\{\tilde{\phi}_j\}$  and  $\{\tilde{\psi}_j\}$ ,  $j = 0, 1, \dots$ , respectively. Notice that the normalization condition for  $\tilde{\sigma}$  implies  $\tilde{\phi}_0 = \tilde{\psi}_0 = 1$ . By Baxter's Theorem [25] we have  $\tilde{\gamma}_j \in \ell^1$  (in fact, the decay is much stronger but  $\ell^1$  is enough for our purposes). Then, let us consider the probability measure  $\sigma$  which has the following Schur parameters

$$\gamma_0, \dots, \gamma_{n-1}, \tilde{\gamma}_0, \tilde{\gamma}_1, \dots$$

We will show that this measure satisfies the Steklov's condition. Denote

$$\gamma_n = \tilde{\gamma}_0, \gamma_{n+1} = \tilde{\gamma}_1, \dots \quad (3.8)$$

The Baxter Theorem implies that  $\sigma$  is purely a.c.,  $\sigma'$  belongs to Wiener's class  $W(\mathbb{T})$ , and  $\sigma'$  is positive on  $\mathbb{T}$ . The first  $n$  orthonormal polynomials corresponding to the measure  $\sigma$  will be  $\{\phi_j\}$ ,  $j = 0, \dots, n-1$ . Let us compute the polynomials  $\phi_j$  and  $\psi_j$ , orthonormal with respect to  $\sigma$ , for the indexes  $j > n$ . Since the second kind polynomials correspond to the Schur parameters  $\{-\gamma_j\}$  (see (3.3)), the recursion can be rewritten in the following matrix form

$$\begin{pmatrix} \phi_{n+m} & \psi_{n+m} \\ \phi_{n+m}^* & -\psi_{n+m}^* \end{pmatrix} = \begin{pmatrix} \mathcal{A}_m & \mathcal{B}_m \\ \mathcal{C}_m & \mathcal{D}_m \end{pmatrix} \begin{pmatrix} \phi_n & \psi_n \\ \phi_n^* & -\psi_n^* \end{pmatrix} \quad (3.9)$$

where  $\mathcal{A}_m, \mathcal{B}_m, \mathcal{C}_m, \mathcal{D}_m$  satisfy

$$\begin{pmatrix} \mathcal{A}_0 & \mathcal{B}_0 \\ \mathcal{C}_0 & \mathcal{D}_0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$\begin{pmatrix} \mathcal{A}_m & \mathcal{B}_m \\ \mathcal{C}_m & \mathcal{D}_m \end{pmatrix} = \frac{1}{\tilde{\rho}_0 \cdots \tilde{\rho}_{m-1}} \begin{pmatrix} z & -\tilde{\gamma}_{m-1} \\ -z\tilde{\gamma}_{m-1} & 1 \end{pmatrix} \cdots \begin{pmatrix} z & -\tilde{\gamma}_0 \\ -z\tilde{\gamma}_0 & 1 \end{pmatrix}$$

and thus depend only on  $\gamma_n, \dots, \gamma_{n+m-1}$  (i.e.,  $\tilde{\gamma}_0, \dots, \tilde{\gamma}_{m-1}$  by (3.8).) Moreover, we have

$$\begin{pmatrix} \tilde{\phi}_m & \tilde{\psi}_m \\ \tilde{\phi}_m^* & -\tilde{\psi}_m^* \end{pmatrix} = \begin{pmatrix} \mathcal{A}_m & \mathcal{B}_m \\ \mathcal{C}_m & \mathcal{D}_m \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Thus,  $\mathcal{A}_m = (\tilde{\phi}_m + \tilde{\psi}_m)/2$ ,  $\mathcal{B}_m = (\tilde{\phi}_m - \tilde{\psi}_m)/2$ ,  $\mathcal{C}_m = (\tilde{\phi}_m^* - \tilde{\psi}_m^*)/2$ ,  $\mathcal{D}_m = (\tilde{\phi}_m^* + \tilde{\psi}_m^*)/2$  and their substitution into (3.9) yields

$$2\phi_{n+m}^* = \phi_n(\tilde{\phi}_m^* - \tilde{\psi}_m^*) + \phi_n^*(\tilde{\phi}_m^* + \tilde{\psi}_m^*) = \tilde{\phi}_m^* \left( \phi_n + \phi_n^* + \tilde{F}_m(\phi_n^* - \phi_n) \right) \quad (3.10)$$

where

$$\tilde{F}_m(z) = \frac{\tilde{\psi}_m^*(z)}{\tilde{\phi}_m^*(z)}.$$

Since  $\{\tilde{\gamma}_n\} \in \ell^1$  and  $\{\gamma_n\} \in \ell^1$ , we have ([25], p. 225)

$$\tilde{F}_m \rightarrow \tilde{F} \text{ as } m \rightarrow \infty \text{ and } \phi_n^* \rightarrow \Pi, \tilde{\phi}_n^* \rightarrow \tilde{\Pi} \text{ as } n \rightarrow \infty.$$

uniformly on  $\overline{\mathbb{D}}$ . The functions  $\Pi$  and  $\tilde{\Pi}$  are the Szegő functions of  $\sigma$  and  $\tilde{\sigma}$ , respectively, i.e., they are the outer functions in  $\mathbb{D}$  that give the factorizations

$$|\Pi|^{-2} = 2\pi\sigma', \quad |\tilde{\Pi}|^{-2} = 2\pi\tilde{\sigma}'. \quad (3.11)$$

In (3.10), send  $m \rightarrow \infty$  to get

$$2\Pi = \tilde{\Pi} \left( \phi_n + \phi_n^* + \tilde{F}(\phi_n^* - \phi_n) \right). \quad (3.12)$$

Thus, the first formula in (3.11) shows that for the sufficiently regular measures, Steklov's condition  $\sigma' \geq \delta/(2\pi)$  is equivalent to

$$\left| \tilde{\Pi} \left( \phi_n + \phi_n^* + \tilde{F}(\phi_n^* - \phi_n) \right) \right| \leq \frac{2}{\sqrt{\delta}}, \quad z = e^{i\theta} \in \mathbb{T}. \quad (3.13)$$

Since  $|\phi_n| = |\phi_n^*|$  on  $\mathbb{T}$ , we have

$$\left| \tilde{\Pi} \left( \phi_n + \phi_n^* + \tilde{F}(\phi_n^* - \phi_n) \right) \right| \leq 2|\tilde{\Pi}| \left( |\phi_n^*| + |\tilde{F}(\phi_n^* - \phi_n)| \right) < 2C_1(\delta)|\tilde{\Pi}| \left( \operatorname{Re} \tilde{F} \right)^{1/2} = 2C_1(\delta)$$

due to (3.6), (3.7), and the second formula in (3.11). Thus, to guarantee (3.13), we only need to take  $C_1(\delta) = \delta^{-1/2}$  in (3.6). In this section, we assume  $\delta$  to be fixed so the exact formulas for  $C(\delta)$  and  $C_1(\delta)$  will not be needed. □

*Proof. (of Theorem 3.1).* The proof will be based on the Decoupling Lemma and will contain two parts. In the first one, we will make the choice for  $\phi_n^*$ ,  $\tilde{F}$  and study their basic properties to check conditions (1),(3)-(5) of the Decoupling Lemma. In the second part, we will verify the normalization condition, i.e., condition (2).

**3.3. The choice of parameters.** In what follows, we take  $\varepsilon_n = n^{-1}$ .

**1. The choice of  $\tilde{F}$ .** Consider two parameters:  $\alpha \in (1/2, 1)$  and  $\rho \in (0, \rho_0)$  where  $\rho_0$  is sufficiently small. Let us emphasize that these parameters are fixed and will not be changed in the estimates below, however many constants in these inequalities will actually depend on them. We do not trace this dependence here.

Take

$$\tilde{F}(z) = \tilde{C}_n (\rho(1 + \varepsilon_n - z)^{-1} + (1 + \varepsilon_n - z)^{-\alpha}), \quad (3.14)$$

where the positive normalization constant  $\tilde{C}_n$  will be chosen later. We have two terms inside the brackets. The first one gives the right growth at point  $z = 1$ :  $(1 + \varepsilon_n - 1)^{-1} = n$  and this choice is motivated by conditions (3) and (5) of the Decoupling Lemma. The role of the second term will be explained later.

We will need more information on  $\tilde{F}$ . Clearly  $\tilde{F}$  is smooth and has a positive real part in  $\overline{\mathbb{D}}$ . Notice that for  $z = e^{i\theta} \in \mathbb{T}$  and  $|\theta| \ll 1$ , we have

$$\begin{aligned} 1 + \varepsilon_n - z &= \varepsilon_n + \frac{\theta^2}{2} - i\theta + O(|\theta|^3), \quad |1 + \varepsilon_n - z|^2 = (\varepsilon_n + \frac{\theta^2}{2} + O(|\theta|^3))^2 + (-\theta + O(|\theta|^3))^2 = \\ &= \varepsilon_n^2 + \theta^2 + \varepsilon_n \theta^2 + O(|\theta|^3) \end{aligned}$$

Since  $w^{-1} = \frac{\overline{w}}{|w|^2}$ , we obtain

$$(1 + \varepsilon_n - z)^{-1} = \frac{\varepsilon_n + \frac{\theta^2}{2} + i\theta + O(|\theta|^3)}{\varepsilon_n^2 + \theta^2 + \varepsilon_n \theta^2 + O(|\theta|^3)} = \frac{\varepsilon_n + i\theta}{\varepsilon_n^2 + \theta^2} + O\left(\frac{\theta^2}{\varepsilon_n^2 + \theta^2}\right). \quad (3.15)$$

The last equality can be verified directly by subtraction. So,

$$\tilde{F}(e^{i\theta}) = \tilde{C}_n \left( \frac{\rho \varepsilon_n}{\varepsilon_n^2 + \theta^2} + i \frac{\rho \theta}{\varepsilon_n^2 + \theta^2} + (1 + \varepsilon_n - z)^{-\alpha} + O\left(\frac{\rho \theta^2}{\varepsilon_n^2 + \theta^2}\right) \right). \quad (3.16)$$

We have

$$(1 + \varepsilon_n - e^{i\theta})^{-\alpha} = ((1 + \varepsilon_n - \cos \theta)^2 + \sin^2 \theta)^{-\alpha/2} \exp(-i\alpha \Gamma_n(\theta)). \quad (3.17)$$

$$\Gamma_n(\theta) = -\arctan\left(\frac{\sin \theta}{1 + \varepsilon_n - \cos \theta}\right) \in (-\pi/2, \pi/2). \quad (3.18)$$

The following bound is true

$$\left\| \frac{\rho \varepsilon_n}{\varepsilon_n^2 + \theta^2} + |(1 + \varepsilon_n - e^{i\theta})^{-\alpha}| + O\left(\frac{\rho \theta^2}{\varepsilon_n^2 + \theta^2}\right) \right\|_{L^1[-\pi, \pi]} \sim 1$$

uniformly in  $n$ . Then, for every fixed  $v > 0$ , we have

$$\rho(1 + \varepsilon_n - z)^{-1} + (1 + \varepsilon_n - z)^{-\alpha} \rightarrow \rho(1 - z)^{-1} + (1 - z)^{-\alpha}, \quad n \rightarrow \infty \quad (3.19)$$

uniformly in  $\{z = e^{i\theta}, |\theta| > v\}$ . Since

$$\int_{|\theta| > v} \operatorname{Re} \left( \rho(1 - e^{i\theta})^{-1} + (1 - e^{i\theta})^{-\alpha} \right) d\theta \sim 1$$

we can choose  $\tilde{C}_n$  to guarantee (3.5) and then  $\tilde{C}_n \sim 1$  uniformly in  $n$ .

Consider the formulas (3.15) and (3.16). They yield

$$|\tilde{F}| \sim \varepsilon_n^{-1} = n \quad \text{for } |\theta| < \varepsilon_n \quad (3.20)$$

and

$$|\tilde{F}| \sim |\theta|^{-1} \quad \text{for } |\theta| > \varepsilon_n. \quad (3.21)$$

Indeed, in the last inequality, the upper bound

$$|\tilde{F}| \lesssim |\theta|^{-1}$$



is immediate. For the lower bound,

$$|\tilde{F}| \geq |\operatorname{Im} \tilde{F}| = \left| \frac{\rho \sin \theta}{(1 + \varepsilon_n - \cos \theta)^2 + \sin^2 \theta} + ((1 + \varepsilon_n - \cos \theta)^2 + \sin^2 \theta)^{-\alpha/2} \sin \left( \alpha \arctan \frac{\sin \theta}{1 + \varepsilon_n - \cos \theta} \right) \right|.$$

These terms have the same signs, so

$$|\operatorname{Im} \tilde{F}| \geq \left| \frac{\rho \sin \theta}{(1 + \varepsilon_n - \cos \theta)^2 + \sin^2 \theta} \right| = \left| \frac{\rho \sin \theta}{2(1 + \varepsilon_n)(1 - \cos \theta) + \varepsilon_n^2} \right| \gtrsim |\theta|^{-1}, \quad \varepsilon_n < |\theta| < \pi/2.$$

For  $\theta : |\theta| > \pi/2$ , the estimate

$$|\tilde{F}| \sim |\theta|^{-1}$$

is a trivial corollary of (3.19).

**2. The choice of  $\phi_n^*$ .** Let  $\phi_n^*$  be chosen as follows

$$\phi_n^*(z) = C_n f_n(z), \quad f_n(z) = P_m(z) + Q_m(z) + Q_m^*(z) \quad (3.22)$$

where  $P_m$  and  $Q_m$  are certain polynomials of degree

$$m = [\delta_1 n] \quad (3.23)$$

where  $\delta_1$  is small and will be specified later. Notice here that  $Q_m^*$  is defined by applying the  $n$ -th order star operation. The constant  $C_n$  will be chosen in such a way that

$$\int_{-\pi}^{\pi} |\phi_n^*|^{-2} d\theta = 2\pi$$

(i.e., (3.4) is satisfied). To prove the Theorem, we only need to show that

$$C_n = \left( \int_{-\pi}^{\pi} |f_n|^{-2} d\theta \right)^{1/2} \sim 1 \quad (3.24)$$

uniformly in  $n$  and that  $f_n$  satisfies the other conditions of the Decoupling Lemma.

The choice of  $\phi_n^*$  is motivated by the following observation. The estimate (3.6) requires

$$|\tilde{F}(z)(\phi_n(z) - \phi_n^*(z))| < C_1(\delta) \left( \operatorname{Re} \tilde{F}(z) \right)^{1/2}, \quad z \in \mathbb{T}.$$

Since  $|F(z)|$  is much larger than  $\operatorname{Re} \tilde{F}(z)$  around  $z = 1$ , the point of growth, the factor  $\phi_n - \phi_n^*$  should provide some cancelation. The sum of the second and the third terms in (3.22), the polynomial  $Q_m + Q_m^*$  has degree  $n$  and is symmetric so it drops out in  $\phi_n - \phi_n^*$ . However, it has zeroes on  $\mathbb{T}$  due to Lemma 3.1 and thus can not be a good choice for  $\phi_n^*$  due to violation of conditions (1) and (2) in Decoupling Lemma.  $P_m$ , the first term in (3.22), will be chosen to achieve a certain balance. It will be small around  $z = 1$  and it will push the zeroes of  $Q_m + Q_m^*$  away from  $\mathbb{D}$  to guarantee (3.4).

Consider the Fejer kernel

$$\mathcal{F}_m(\theta) = \frac{1}{m} \frac{\sin^2(m\theta/2)}{\sin^2(\theta/2)} = \frac{1}{m} \frac{1 - \cos(m\theta)}{1 - \cos(\theta)}, \quad \mathcal{F}_m(0) = m \quad (3.25)$$

and the Taylor approximation to the function  $(1 - z)^{-\alpha}$ , i.e.,

$$R_{(k, \alpha)}(z) = 1 + \sum_{j=1}^k d_j z^j$$

(see Appendixes for the detailed discussion). We define  $Q_m$  as an analytic polynomial without zeroes in  $\mathbb{D}$  which gives Fejer-Riesz factorization

$$|Q_m(z)|^2 = \mathcal{G}_m(\theta) + |R_{(m,\alpha/2)}(e^{i\theta})|^2. \quad (3.26)$$

$$\mathcal{G}_m(\theta) = \mathcal{F}_m(\theta) + \frac{1}{2}\mathcal{F}_m\left(\theta - \frac{\pi}{m}\right) + \frac{1}{2}\mathcal{F}_m\left(\theta + \frac{\pi}{m}\right). \quad (3.27)$$

Clearly, the right hand side of (3.26) is a positive trigonometric polynomial of degree  $m$  so this factorization is possible and  $Q_m$  is unique up to a unimodular factor. We choose this factor in such a way that  $Q_m(0) > 0$ . Since  $Q_m$  is an outer function, we have the following canonical representation (see [9], page 24)

$$Q_m(z) = \exp\left(\frac{1}{2\pi} \int_{-\pi}^{\pi} C(z, e^{i\theta}) \ln |Q_m(e^{i\theta})| d\theta\right), \quad |z| < 1. \quad (3.28)$$

Notice that  $Q_m^*$  is a polynomial of degree  $n$  with positive leading coefficient. Since  $|Q_m(e^{i\theta})|$  is even in  $\theta$ , this representation shows that  $\mathcal{H}(z) = \ln Q_m(z)$  is analytic in  $\mathbb{D}$  and has real Taylor coefficients (indeed,  $\overline{\mathcal{H}(z)} = \mathcal{H}(\bar{z})$ ). That, on the other hand, implies that  $Q_m(z) = e^{\mathcal{H}(z)}$  has real coefficients as well.

For  $P_m$ , we take

$$P_m(z) = Q_m(z)(1-z)(1-0.1R_{(m, -(1-\alpha))}(z)) \quad (3.29)$$

and  $\deg P_m = 2m + 1 < n$  by the choice of small  $\delta_1$ . Consequently,  $\deg \phi_n^* = n$ .

Now that we have chosen  $\tilde{F}$  and  $\phi_n^*$ , it is left to show that they satisfy the conditions of the Decoupling Lemma. The second term in (3.14) and the structure of (3.29) will become important in what follows.

**1.  $f_n$  has no zeroes in  $\overline{\mathbb{D}}$ .** For  $f_n$ , we can write

$$f_n = Q_m(z) \left( (1-z)(1-0.1R_{(m, -(1-\alpha))}(z)) + 1 + z^n e^{-2i\phi} \right), \quad z \in \mathbb{T} \quad (3.30)$$

where

$$e^{-2i\phi(\theta)} = \frac{\overline{Q_m(e^{i\theta})}}{Q_m(e^{i\theta})}$$

so  $\phi$  is an argument of  $Q_m$ . The polynomial  $Q_m$  has no zeroes in  $\mathbb{D}$  and

$$(1-z)(1-0.1R_{(m, -(1-\alpha))}(z)) + 1 + \frac{Q_m^*}{Q_m} \quad (3.31)$$

is analytic in  $\mathbb{D}$  and has positive real part. Indeed,

$$\operatorname{Re} \left( 1 + \frac{Q_m^*}{Q_m} \right) \geq 0, \quad z \in \mathbb{T}$$

since  $|Q_m| = |Q_m^*|$ .

Since  $Q_m$  has real coefficients,  $Q_m(1)$  is real. Furthermore, since  $Q_m(0) > 0$ ,  $Q_m$  is real at the real line and  $Q_m \neq 0$  in  $\mathbb{D}$ , then  $Q_m(1) > 0$ . So  $\phi(0) = 0$  and

$$\operatorname{Re} \left( 1 + \frac{Q_m^*}{Q_m} \right) = 2, \quad z = 1.$$

For the first term in (3.31), we have

$$\operatorname{Re} [(1-z)(1-0.1R_{(m, -(1-\alpha))}(z))] = (1 - \cos \theta)(1 - 0.1X) - 0.1Y \sin \theta \quad (3.32)$$

where

$$X = \operatorname{Re} R_{(m, -(1-\alpha))}, \quad Y = \operatorname{Im} R_{(m, -(1-\alpha))}.$$

Notice that

$$|R_{(m, -(1-\alpha))}| < 2 \quad \text{for } z \in \mathbb{D}. \quad (3.33)$$

So  $|X| < 2$  in  $\mathbb{D}$  as well. The function  $Y$  is odd in  $\theta$  and  $Y(\theta) < 0$  for  $\theta \in (0, \pi)$  (as follows from the Lemma 8.5 in Appendix A for small  $\theta$ ; see [4], Theorem 7.3.5 for the general case). Thus, the function  $f_n(z)$  has a positive real part on  $\mathbb{T}$  and, in particular, has no zeroes in  $\mathbb{D}$ . We conclude then that  $\phi_n^*$  has no zeroes in  $\mathbb{D}$ .

**2. The growth at  $z = 1$ .** The formula (3.30) implies

$$f_n(1) = 2Q_m(1).$$

Then, (3.25) and (3.26) yield

$$|f_n(1)| \gtrsim \sqrt{m} \sim \sqrt{n}$$

due to (3.23).

**3. Steklov's condition.** We need to check (3.6) with  $\phi_n$  replaced by  $f_n$ . From (3.26) and (3.30), we get

$$|f_n|^2 \lesssim |Q_m|^2 = \mathcal{G}_m(\theta) + |R_{(m, \alpha/2)}(e^{i\theta})|^2.$$

Lemma 8.3 implies

$$|R_{(m, \alpha/2)}(e^{i\theta})|^2 \lesssim (\varepsilon_m + |\theta|)^{-\alpha} \lesssim (\varepsilon_n + |\theta|)^{-\alpha}.$$

The exact form of the Fejer's kernel (3.25) gives

$$\mathcal{G}_m(\theta) \lesssim \frac{m}{m^2\theta^2 + 1} \lesssim \frac{n}{n^2\theta^2 + 1} = \frac{\varepsilon_n}{\varepsilon_n^2 + \theta^2}.$$

We have

$$\begin{aligned} \operatorname{Re} \tilde{F} &= \frac{\rho(1 + \varepsilon_n - \cos \theta)}{(1 + \varepsilon_n - \cos \theta)^2 + \sin^2 \theta} \\ &+ \left( (1 + \varepsilon_n - \cos \theta)^2 + \sin^2 \theta \right)^{-\alpha/2} \cos \left( \alpha \arctan \frac{\sin \theta}{1 + \varepsilon_n - \cos \theta} \right). \end{aligned}$$

Since  $\alpha < 1$  and  $|\arctan(\cdot)| < \pi/2$ , we get

$$\operatorname{Re} \tilde{F} \sim \frac{\rho(\varepsilon_n + \theta^2)}{\varepsilon_n^2 + \theta^2} + \frac{1}{(\varepsilon_n^2 + \theta^2)^{\alpha/2}}.$$

The estimate

$$\frac{\theta^2}{\varepsilon_n^2 + \theta^2} \lesssim \frac{1}{(\varepsilon_n^2 + \theta^2)^{\alpha/2}},$$

uniform in  $\theta$ , implies

$$\operatorname{Re} \tilde{F} \sim \frac{\varepsilon_n}{\varepsilon_n^2 + \theta^2} + (\varepsilon_n^2 + \theta^2)^{-\alpha/2}.$$

Therefore,

$$|f_n| \lesssim |Q_m| \lesssim (\operatorname{Re} \tilde{F})^{1/2}. \quad (3.34)$$

Then, for the second term in (3.6), we get

$$|\tilde{F}(f_n - f_n^*)|^2 = |\tilde{F}(P_m - P_m^*)|^2 \lesssim |\tilde{F}(1 - z)|^2 |Q_m|^2.$$

The uniform bounds

$$|\tilde{F}(1 - z)| \lesssim 1, \quad |Q_m|^2 \lesssim \operatorname{Re} \tilde{F}$$

together with (3.34), imply (3.6) with  $\phi_n$  replaced by  $f_n$ .

**3.4. Normalization: checking condition (2) of the Decoupling Lemma.** We only need to check now that

$$\int_{\mathbb{T}} \frac{1}{|f_n|^2} d\theta \sim 1$$

(see (3.24)). It is sufficient to consider  $\theta \in [0, \pi]$  as  $|f_n(e^{i\theta})|$  is even.

Our first goal is to obtain a convenient lower bound on  $|f_n|^2$ . Notice that (3.26) yields

$$|Q_m(e^{i\theta})|^2 \geq |R_{(m, \alpha/2)}(e^{i\theta})|^2$$

and the Lemma 8.3 from Appendix A gives (we should use notation  $\beta = 1 - \alpha > 0$  here)

$$|f_n|^2 \gtrsim (m^{-1} + |\theta|)^{-\alpha} \left| (1 - z)(1 - 0.1R_{(m, -\beta)}(z)) + 1 + z^n e^{2i\phi} \right|^2.$$

Then, the representation (3.32) leads to

$$\begin{aligned} & \left| (1 - z)(1 - 0.1R_{(m, -\beta)}(z)) + 1 + z^n e^{2i\phi} \right|^2 = \\ & \left( (1 - \cos \theta)(1 - 0.1X) - 0.1Y \sin \theta + 1 + \cos(n\theta - 2\phi) \right)^2 + \\ & \left( -0.1Y(1 - \cos \theta) - (1 - 0.1X) \sin \theta + \sin(n\theta - 2\phi) \right)^2. \end{aligned} \quad (3.35)$$

For the first term, we have

$$(1 - \cos \theta)(1 - 0.1X) - 0.1Y \sin \theta + 1 + \cos(n\theta - 2\phi) \geq -0.1Y \sin \theta \gtrsim \theta^{2-\alpha} \quad (3.36)$$

where the last inequality follows from Lemma 8.5. Thus,

$$|f_n|^2 \gtrsim (m^{-1} + |\theta|)^{-\alpha} \left[ (\theta^{2-\alpha})^2 + (\Psi(\theta) + \sin(n\theta - 2\phi))^2 \right] \quad (3.37)$$

where

$$\Psi = \text{Im} \left( (1 - z)(1 - 0.1R_{(m, -\beta)}(z)) \right) = -0.1Y(1 - \cos \theta) - (1 - 0.1X) \sin \theta.$$

In what follows, we will control  $\Psi$  and  $\phi$  to analyze  $\Psi(\theta) + \sin(n\theta - 2\phi)$  in (3.37). We will locate the zeroes  $\{\theta_j\}$  of this highly oscillatory function and will show that away from these points the normalization condition is easily satisfied. More delicate analysis will be needed to integrate  $|f_n|^{-2}$  over small neighborhoods of  $\{\theta_j\}$ .

To bound  $|\Psi'(\theta)|$ , we use Lemma 8.6 and (3.33). Indeed,

$$\begin{aligned} |\Psi'(\theta)| &= \left| \text{Im} \frac{d}{d\theta} \left[ (1 - z)(1 - 0.1R_{(m, -\beta)}(z)) \right] \right| \leq \\ &\leq \left| (1 - 0.1R_{(m, -\beta)}(e^{i\theta})) \right| + 0.1 \left| (1 - e^{i\theta}) \frac{d}{d\theta} R_{(m, -\beta)}(e^{i\theta}) \right| \leq \\ &\leq (1 + 0.1 \cdot 2) + 0.1 \cdot 2|\theta| \cdot O \left( \max \left( \frac{1}{n}, |\theta| \right)^{\beta-1} \right) = O(1). \end{aligned}$$

Then,

$$|\Psi'(\theta)| \lesssim 1 \quad (3.38)$$

uniformly in  $n$ . This estimate and  $\Psi(0) = 0$  imply

$$|\Psi(\theta)| \lesssim |\theta| \quad (3.39)$$

by integration.

For the phase  $\phi$ , we have  $\phi(0) = 0$  and

$$|\phi'(\theta)| \lesssim m$$

where the last inequality is proved in Appendix B. Since we have the derivative of  $\phi$  under control,

$$(n\theta - 2\phi(\theta))' = n + O(\delta_1 n) . \quad (3.40)$$

By making  $\delta_1$  small, we can make sure that the function  $n\theta - 2\phi(\theta)$  is monotonically increasing and

$$n/2 < (n\theta - 2\phi(\theta))' < 2n . \quad (3.41)$$

To study the zeroes  $\{\theta_j\}$ , we first introduce auxiliary points  $\{\widehat{\theta}_j\}$ . The monotonicity of  $n\theta - 2\phi(\theta)$  allows us to uniquely define  $\{\widehat{\theta}_j\}$  as solutions to the equation:

$$n\widehat{\theta}_j - 2\phi(\widehat{\theta}_j) = \pi(j - \frac{1}{2}), \quad j = 0, \dots, [cn] . \quad (3.42)$$

Then,  $\sin(n\widehat{\theta}_j - 2\phi(\widehat{\theta}_j)) = (-1)^{j-1}$ . On the other hand, from (3.41), one has

$$\frac{(2j-1)\pi}{4n} < \widehat{\theta}_j < \frac{2\pi j}{n} , \quad (3.43)$$

By the estimate (3.39), we can choose some small positive constant  $c > 0$  so that for  $j = 0, 1, \dots, [cn]$  the expression  $\Psi(\theta) + \sin(n\theta - 2\phi(\theta))$  changes the sign on  $[\widehat{\theta}_j, \widehat{\theta}_{j+1}]$ . Indeed,

$$\begin{aligned} (-1)^j \left[ \Psi(\widehat{\theta}_j) + \sin(n\widehat{\theta}_j - 2\phi(\widehat{\theta}_j)) \right] &= (-1)^j \left[ \Psi(\widehat{\theta}_j) + (-1)^{j-1} \right] < |\Psi(\widehat{\theta}_j)| - 1 = O\left(\frac{2\pi j}{n}\right) - 1 < 0 ; \\ (-1)^j \left[ \Psi(\widehat{\theta}_{j+1}) + \sin(n\widehat{\theta}_{j+1} - 2\phi(\widehat{\theta}_{j+1})) \right] &> 1 - |\Psi(\widehat{\theta}_{j+1})| = 1 - O\left(\frac{2\pi j}{n}\right) > 0 . \end{aligned}$$

Moreover, we require that  $c$  is chosen such that

$$|\Psi(\widehat{\theta}_j)| < \frac{1}{\sqrt{2}}, \quad j = 0, 1, \dots, [cn]$$

and that  $v$ , defined as

$$v = \widehat{\theta}_{[cn]},$$

is smaller than the parameter  $v$  from the Lemmas 8.3, 8.4, 8.5 in the Appendix A. Now, let us show that there is the unique point  $\theta_j$  such that

$$\theta_j \in [\widehat{\theta}_j, \widehat{\theta}_{j+1}] : \quad \Psi(\theta_j) + \sin(n\theta_j - 2\phi(\theta_j)) = 0 . \quad (3.44)$$

The existence of such  $\theta_j$  is a simple corollary of continuity and sign change.

Note, that the function  $\Xi(\theta) = n\theta - 2\phi(\theta) - \pi j$ , restricted to the segment  $[\widehat{\theta}_j, \widehat{\theta}_{j+1}]$ , takes values from  $[-\frac{\pi}{2}, \frac{\pi}{2}]$ . So, for each  $\theta_j$  that satisfies (3.44), we have  $\sin(\Xi(\theta_j)) = (-1)^{j-1} \Psi(\theta_j)$  and

$$n\theta_j - 2\phi(\theta_j) + (-1)^j \arcsin(\Psi(\theta_j)) = \pi j . \quad (3.45)$$

If, for fixed  $j$ , there are several solutions  $\theta_j$  to (3.44), then the derivative of the function

$$n\theta - 2\phi(\theta) + (-1)^j \arcsin(\Psi(\theta))$$

in the left hand side of (3.45) is non-positive at at least one of these  $\theta_j$ . However the lower estimate on the derivative reads

$$n - 2\phi'(\theta_j) + \frac{(-1)^j \Psi'(\theta_j)}{\sqrt{1 - \Psi^2(\theta_j)}} > n - O(m) - \frac{O(1)}{\sqrt{1 - \frac{1}{2}}} > \frac{n}{2}$$

which shows that  $\theta_j$  is unique.

Since  $|\Psi(\widehat{\theta}_j)| < \frac{1}{\sqrt{2}}$ , we have  $\arcsin|\Psi(\widehat{\theta}_j)| < \frac{\pi}{4}$ . Now, from (3.40), (3.42) and (3.45), one gets

$$\theta_j - \widehat{\theta}_j > \frac{1}{2n} \left( \frac{\pi}{2} - \frac{\pi}{4} \right) , \quad \widehat{\theta}_{j+1} - \theta_j > \frac{1}{2n} \left( \frac{\pi}{2} - \frac{\pi}{4} \right) , \quad \text{i.e.,} \quad I_j = \left[ \theta_j - \frac{0.01}{n}, \theta_j + \frac{0.01}{n} \right] \subset [\widehat{\theta}_j, \widehat{\theta}_{j+1}] .$$

If  $\theta \in [\widehat{\theta}_j, \widehat{\theta}_{j+1}] \setminus I_j$ , then

$$|(n\theta - 2\phi(\theta)) - (n\theta_j - 2\phi(\theta_j))| > \frac{n}{2} \cdot \frac{0.01}{n} > \frac{1}{200}. \quad (3.46)$$

The estimate for the integral over  $I_0 = [0, 0.01n^{-1}]$  is easy since on that arc we have  $|Q_m| \sim m$  (by (3.26)) and

$$|f_n|^2 \gtrsim m(1 + \cos(n\theta - 2\phi(\theta)))^2$$

by (3.30), (3.36). The bound on the derivative of  $\phi$  implies that  $|f_n| \sim \sqrt{m}$  on  $I_0$  so

$$\int_{I_0} |f_n|^{-2} d\theta \sim n^{-2}.$$

Let  $\theta \in I_j$  and assume that some  $\xi$  is located between  $\theta_j$  and  $\theta$ . From the definition of  $I_j$ , we get  $|\xi - \theta_j| < \frac{1}{100n}$ . Therefore,

$$\begin{aligned} |n\xi - 2\phi(\xi) - \pi j| &\leq |n\theta_j - 2\phi(\theta_j) - \pi j| + |(n\xi - 2\phi(\xi)) - (n\theta_j - 2\phi(\theta_j))| \leq \\ &\leq |\arcsin \Psi(\theta_j)| + 2n|\xi - \theta_j| \leq \frac{\pi}{4} + \frac{2}{100}; \end{aligned}$$

and

$$|\cos(n\xi - 2\phi(\xi))| \geq \cos\left(\frac{\pi}{4} + \frac{2}{100}\right) > \frac{1}{2}. \quad (3.47)$$

So, for  $\theta \in I_j$ ,  $j \geq 1$ , we have

$$|\Psi(\theta) + \sin(n\theta - 2\phi(\theta))| = \left| \int_{\theta_j}^{\theta} \left( \Psi'(\xi) + \cos(n\xi - 2\phi(\xi))(n - 2\phi'(\xi)) \right) d\xi \right| \sim n|\theta - \theta_j|.$$

For  $\theta \in [\widehat{\theta}_j, \widehat{\theta}_{j+1}] \setminus I_j$ ,  $j \geq 1$ , one gets

$$|\sin(n\theta - 2\phi(\theta)) + \Psi(\theta)| \geq |\sin(n\theta - 2\phi(\theta)) - \sin(n\theta_j - 2\phi(\theta_j))| - |\Psi(\theta) - \Psi(\theta_j)|$$

by the triangle inequality. Then, for the last term, we apply (3.38) to get

$$|\Psi(\theta) - \Psi(\theta_j)| \lesssim |\theta - \theta_j| \leq |\widehat{\theta}_{j+1} - \widehat{\theta}_j| = O(n^{-1}).$$

For the first one,

$$\sin x_1 - \sin x_2 = 2 \sin\left(\frac{x_1 - x_2}{2}\right) \cos\left(\frac{x_1 + x_2}{2}\right)$$

where  $x_1 = n\theta - 2\phi(\theta)$  and  $x_2 = n\theta_j - 2\phi(\theta_j)$ . Next, we use (3.46) to write

$$\frac{\pi}{2} \geq |x_1 - x_2| \geq \frac{1}{200}.$$

Consider  $(x_1 + x_2)/2$ . For  $x_2$ , we have  $x_2 = \pi j - (-1)^j \arcsin \Psi(\theta_j)$  with  $|\arcsin \Psi(\theta_j)| \leq \pi/4$ . Then, for every  $x_1 \in [\pi j - \pi/2, \pi j + \pi/2]$ , we get

$$\left| \cos\left(\frac{x_1 + x_2}{2}\right) \right| > \cos(3\pi/8).$$

Therefore,

$$|\Psi(\theta) + \sin(n\theta - 2\phi(\theta))| \sim 1$$

outside  $\cup_j I_j$ .

Now, let us obtain the estimates outside the small fixed arc  $\{\theta : |\theta| > \nu\}$ . The bound (3.30) implies

$$|Q_m|^2 \left[ \operatorname{Re} \left( (1-z)(1 - 0.1R_{(m, -(1-\alpha))}(z)) \right) \right]^2 \leq |f_n|^2 \leq |Q_m|^2 (2 + |(1-z)(1 - 0.1R_{(m, -(1-\alpha))}(z))|)^2.$$

We have the uniform convergence

$$R_{(m, -(1-\alpha))}(z) \rightarrow (1-z)^{1-\alpha}, \quad m \rightarrow \infty, z \in \mathbb{T}$$

and

$$|Q_m(z)| \rightarrow |1-z|^{-\alpha}, \quad m \rightarrow \infty, \quad z = e^{i\theta}, |\theta| > v.$$

The direct calculation shows (see, e.g., (3.36)) that

$$\operatorname{Re}\left((1-z)(1-0.1R_{(m, -(1-\alpha))}(z))\right) \sim 1, \quad |\theta| > v.$$

Consequently,

$$\int_{|\theta|>v} |f_n|^{-2} d\theta \sim C(v, \alpha). \quad (3.48)$$

From (3.37), we have

$$\begin{aligned} \|f_n^{-1}\|_2^2 &\lesssim \int_{[0,v] \cap (\cup_j I_j)^c} \theta^\alpha d\theta + \sum_{j=1}^{[cn]} \theta_j^\alpha \int_{-0.01n^{-1}}^{0.01n^{-1}} \frac{d\theta}{(\theta_j^{2-\alpha})^2 + n^2 \theta^2} + n^{-2} + \int_{|\theta|>v} |f_n|^{-2} d\theta \lesssim \\ &\lesssim 1 + \frac{1}{n} \sum_{j=1}^{[cn]} \theta_j^{-2+2\alpha} \sim 1 + \int_0^1 \theta^{2\alpha-2} d\theta \lesssim 1 \end{aligned}$$

where we used (3.43), (3.44), and  $\alpha \in (1/2, 1)$ . Together with (3.48), that implies

$$\int_{\mathbb{T}} |f_n|^{-2} d\theta \sim 1$$

and the proof of Theorem 3.1 is finished.  $\square$

**Remark.** It is immediate from the proof that the constructed polynomial satisfies the following bound:

$$|\phi_n(e^{i\theta}, \sigma)|^2 \leq C(\alpha) \left( \frac{n}{1 + (n\theta)^2} + \frac{1}{|\theta|^\alpha} \right), \quad \alpha \in (0.5, 1). \quad (3.49)$$

#### 4. MEASURE OF ORTHOGONALITY

Our method allows one to compute a measure of orthogonality  $\sigma$  for which the orthonormal polynomial has the required size and it is interesting to compare it to the results on the maximizers we obtained before. The calculations given below will show that  $\sigma$  is purely absolutely continuous. Its density can be represented as a sum of background  $\mathcal{B}(\theta)$ ,  $0 < C_1 \leq \mathcal{B}(\theta) \leq C_2$ , and a combination of “peaks” positioned at  $\tilde{\theta}_j$  to be defined later. Qualitatively, each peak resembles the mollification of the point mass by the Poisson kernel. Our analysis can establish the parameters of mollification and a “mass” assigned to each peak.

The formulas (3.11) and (3.12) yield

$$\sigma' = \frac{4\tilde{\sigma}'}{|\phi_n + \phi_n^* + \tilde{F}(\phi_n^* - \phi_n)|^2} = \frac{2 \operatorname{Re} \tilde{F}}{\pi |\phi_n + \phi_n^* + \tilde{F}(\phi_n^* - \phi_n)|^2}.$$

This expression is explicit as we know the formulas for all functions involved. We have

$$\sigma'^{-1} = \frac{C_n |Q_m|^2}{\operatorname{Re} \tilde{F}} \cdot \left| 2(1 + e^{i(n\theta-2\phi)}) + H_n(e^{i\theta}) + e^{i(n\theta-2\phi)} \overline{H_n(e^{i\theta})} + \tilde{F} \left( H_n(e^{i\theta}) - e^{i(n\theta-2\phi)} \overline{H_n(e^{i\theta})} \right) \right|^2 \quad (4.1)$$

where

$$H_n(z) = (1-z)(1-0.1R_{(m, -(1-\alpha))}(z)), \quad C_n \sim 1 \quad (4.2)$$

as follows from (3.24), (3.29).

Consider the first factor. We can apply (3.20), (3.21), and Lemma 8.3 to get

$$\frac{|Q_m|^2}{\operatorname{Re} \tilde{F}} \sim 1.$$

Recall that  $\tilde{F}$  is given by

$$\tilde{F}(z) = \tilde{C}_n(\rho(1 + \varepsilon_n - z)^{-1} + (1 + \varepsilon_n - z)^{-\alpha}), \quad \tilde{C}_n = (\rho/(1 + \varepsilon_n) + (1 + \varepsilon_n)^{-\alpha})^{-1} \quad (4.3)$$

where the last formula for  $\tilde{C}_n$  comes from the normalization (3.5) and

$$\frac{1}{2\pi} \int_{\mathbb{T}} \operatorname{Re} \tilde{F}(e^{i\theta}) d\theta = \operatorname{Re} \tilde{F}(0)$$

(the Mean Value Formula for a harmonic function continuous in  $\mathbb{D}$ ).

Substitution into the second factor in (4.1) gives

$$(\sigma')^{-1} \sim \left| (2 + \overline{H}_n(1 - \tilde{F})) \left( e^{i(n\theta - 2\phi)} + \frac{2 + H_n(1 + \tilde{F})}{2 + \overline{H}_n(1 - \tilde{F})} \right) \right|^2. \quad (4.4)$$

For  $z = e^{i\theta}$  and small positive  $\theta$  (the negative values can be handled similarly), we have

$$H_n(e^{i\theta}) = -i\theta + 0.1\theta i R_{(m, -(1-\alpha))}(e^{i\theta}) + O(|\theta|^2)$$

and Lemma 8.5 can be used for  $R_{(m, -(1-\alpha))}$ . Next, consider  $\tilde{F}$ . It can be written as

$$\tilde{F}(e^{i\theta}) = \tilde{C}_n \left( \frac{\rho}{\varepsilon_n - i\theta} + \frac{1}{(\varepsilon_n - i\theta)^\alpha} \right) + O(1)$$

Therefore, for the first factor in (4.4), we have

$$|2 + \overline{H}_n(1 - \tilde{F})| \sim 1 \quad (4.5)$$

when  $\rho \in (0, \rho_0(\alpha))$  and  $\rho_0(\alpha)$  is small.

Consider

$$J = \frac{2 + H_n(1 + \tilde{F})}{2 + \overline{H}_n(1 - \tilde{F})}.$$

Notice first that  $|J| > 1$  for  $\theta \neq 0$ . Indeed,

$$\begin{aligned} |J|^2 - 1 &= \frac{|2 + H_n(1 + \tilde{F})|^2 - |2 + \overline{H}_n(1 - \tilde{F})|^2}{|2 + \overline{H}_n(1 - \tilde{F})|^2} = \frac{4 \operatorname{Re} \left( (H_n + \overline{H}_n + |H_n|^2) \overline{\tilde{F}} \right)}{|2 + \overline{H}_n(1 - \tilde{F})|^2} \\ &= \frac{4(\operatorname{Re} \tilde{F})(|H_n|^2 + 2 \operatorname{Re} H_n)}{|2 + \overline{H}_n(1 - \tilde{F})|^2} \end{aligned}$$

and the last expression is positive by the choice of  $\tilde{F}$  and  $H_n$ . For  $\theta = 0$ , we have  $J(0) = 1$ .

For small  $\theta$ , the following asymptotics holds

$$|J|^2 - 1 \sim |\theta|^{2-\alpha} \left( \frac{\rho \varepsilon_n}{\varepsilon_n^2 + \theta^2} + |\theta|^{-\alpha} \right), \quad |\theta| > 0.1n^{-1} \quad (4.6)$$

and

$$|J|^2 - 1 \sim \left( \frac{\rho \varepsilon_n}{\varepsilon_n^2 + \theta^2} + \varepsilon_n^{-\alpha} \right) |\theta| n^{\alpha-1}, \quad |\theta| < 0.1n^{-1}.$$

If we write

$$J = r(\theta) e^{i\Upsilon(\theta)}, \quad \Upsilon(0) = 0$$



then

$$(\sigma')^{-1} \sim |e^{i(n\theta-2\phi(\theta)-\Upsilon(\theta))} + r(\theta)|^2 = (\cos(n\theta-2\phi(\theta)-\Upsilon(\theta)) + r(\theta))^2 + \sin^2(n\theta-2\phi(\theta)-\Upsilon(\theta)) . \quad (4.7)$$

Consider  $\{\tilde{\theta}_j\}$ , the solutions to

$$n\theta - 2\phi(\theta) - \Upsilon(\theta) = \pi + 2\pi j, \quad |j| < N_3$$

that belong to some small fixed arc  $|\theta| < v$ . We have  $\Upsilon(0) = 0$  and the direct estimation gives

$$|\Upsilon'(\theta)| < 0.1n$$

uniformly for all  $\theta$ . Indeed, it is sufficient to show that

$$\left| \partial_\theta \left( \frac{2 + H_n(1 + \tilde{F})}{2 + \overline{H}_n(1 + \tilde{F})} \right) \right| < 0.01n \quad (4.8)$$

and

$$\left| \partial_\theta \left( \frac{2 + \overline{H}_n(1 - \tilde{F})}{2 + H_n(1 - \tilde{F})} \right) \right| < 0.01n . \quad (4.9)$$

The both inequalities are proved in Lemma 8.8 from Appendix B.

Now we can argue that the distance between the consecutive  $\{\tilde{\theta}_j\}$  is of size  $n^{-1}$  and

$$\sigma' \sim \frac{1}{(r(\tilde{\theta}_j) - 1)^2 + n^2(\theta - \tilde{\theta}_j)^2} = \tilde{m}_j \frac{\tilde{y}_j}{\tilde{y}_j^2 + (\theta - \tilde{\theta}_j)^2}$$

on  $\theta : |\theta - \tilde{\theta}_j| < Cn^{-1}$ . In the Poisson kernel, the mollification parameter  $\tilde{y}_j$  is

$$\tilde{y}_j = \frac{r(\tilde{\theta}_j) - 1}{n}$$

and the mass  $\tilde{m}_j$  is given by

$$\tilde{m}_j = \frac{1}{n(r(\tilde{\theta}_j) - 1)}$$

Notice, that

$$\sum_{j=-N_3}^{N_3} \tilde{m}_j \lesssim \int_{0.1n^{-1} < |\theta| < v} \frac{d\theta}{r(\theta) - 1} < \infty$$

as follows from (4.6) and  $\alpha \in (1/2, 1)$ . Away from these  $\{\tilde{\theta}_j\}$  the density is  $\sim 1$ .

**Remark.** In the estimates above, we assumed that  $\rho$  is small:  $\rho \in (0, \rho_0)$ . The choice of  $\rho_0$  is made in Lemma 8.8 (see the Remark after it) and in (4.5). Thus, we first fix a parameter  $\alpha$  and then fix  $\rho$ . In fact, we need  $\rho$  to be small only to control  $\sigma'$  and it is irrelevant for the proof of the main Theorem.

The measure  $\sigma$  constructed in the proof has no singular part. Its regularity can be summarized in the following Lemma. For  $\delta \in (0, 1), p \in [1, \infty], C > 0$ , let us introduce the following class of measures given by a weight

$$S_\delta^{(p,C)} = \{\sigma : d\sigma = w(\theta)d\theta, w \geq \delta/(2\pi), \|w\|_1 = 1, \|w\|_p \leq C\}$$

and let

$$M_{n,\delta}^{(p,C)} = \sup_{\sigma \in S_\delta^{(p,C)}} \|\phi_n(z, \sigma)\|_{L^\infty(\mathbb{T})}$$

**Lemma 4.1.** *For every  $p \in [1, \infty)$  there is  $\delta_0 \in (0, 1)$  and  $C > 0$  such that*

$$M_{n, \delta_0}^{(p, C)} \sim \sqrt{n}$$

*Proof.* The upper estimate is obvious since  $M_{n, \delta}^{(p, C)} \leq M_{n, \delta}$ . For the lower one, we only need to show that for every large  $p$  one can take  $\alpha$  in the proof of Theorem 3.1 so close to 1 that  $\|w\|_p < C$  where  $w = \sigma'$  and  $C$  is independent of  $n$ . The estimate (4.7) and the bounds on the derivatives of  $\phi$  and  $\Upsilon$  yield

$$\begin{aligned} \int_{\mathbb{T}} w^p d\theta &\lesssim 1 + \sum_{j=1}^{N_3} \int_0^{Cn^{-1}} \frac{d\theta}{((r(\widehat{\theta}_j) - 1)^2 + n^2 \theta^2)^p} \lesssim 1 + \sum_{j=1}^{N_3} \frac{1}{n(r(\widehat{\theta}_j) - 1)^{2p-1}} \\ &\lesssim 1 + \int_0^1 \frac{d\theta}{|\theta|^{(2p-1)(2-2\alpha)}} < \infty \end{aligned}$$

if  $\alpha \in ((4p-3)/(4p-2), 1)$ . Here we used (4.6) to estimate  $r(\theta) - 1$ .  $\square$

### Part 3. Bernstein's method and localization. The proofs of Theorem 0.3 and Theorem 0.4

In this part, we will use the “localization principle” to first prove the lower bounds on  $M_{n, \delta}$  in the full range of  $\delta$  (Theorem 0.3) and then iterate this construction and prove Theorem 0.4.

#### 5. THE METHOD BY BERNSTEIN AND LOCALIZATION PRINCIPLE.

Given a weight  $w$  on  $[-\pi, \pi]$ , we define

$$\lambda(w) = \exp \left( \frac{1}{4\pi} \int_{\mathbb{T}} \ln(2\pi w(\theta)) d\theta \right), \quad \Lambda(w) = \sqrt{\|w\|_1}.$$

We have

$$\phi_n(z, w) = \|w\|_1^{-1/2} \phi_n(z, w/\|w\|_1)$$

and (0.17), (0.18) yield

$$\exp \left( \frac{1}{4\pi} \int_{\mathbb{T}} \ln(2\pi w(\theta)/\|w\|_1) d\theta \right) \leq \left| \frac{\Phi_n(z, w)}{\phi_n(z, w/\|w\|_1)} \right| \leq 1.$$

So,

$$\lambda(w) \leq \left| \frac{\Phi_n(z, w)}{\phi_n(z, w)} \right| \leq \Lambda(w). \quad (5.1)$$

In [6], S. Bernstein studied the asymptotics of the polynomials when the weight of orthogonality is regular and introduced a method which we will use when proving the following Theorem.

**Theorem 5.1.** *Let  $w_{1(2)}$  be two weights on  $[-\pi, \pi]$  so that*

$$w_1(\theta) = w_2(\theta), \quad \theta \in [-\epsilon, \epsilon]. \quad (5.2)$$

*Then*

$$\left| \frac{\phi_n(1, w_1)}{\phi_n(1, w_2)} \right| \leq \frac{\Lambda(w_2)}{\lambda(w_1)} + \frac{4\Lambda(w_1)}{\epsilon\lambda(w_1)} \left( \int_{|\theta| > \epsilon} |\phi_n(e^{i\theta}, w_1) \phi_n(e^{i\theta}, w_2)| (w_1 + w_2) d\theta \right) \quad (5.3)$$

*for all  $n$ .*

*Proof.* Following Bernstein, we write

$$\Phi_n(z, w_1) = \varrho_n \phi_n(z, w_2) + \sum_{j=0}^{n-1} \phi_j(z, w_2) \int_{-\pi}^{\pi} \Phi_n(e^{i\theta}, w_1) \bar{\phi}_j(e^{i\theta}, w_2) w_2(\theta) d\theta \quad (5.4)$$

with some coefficient  $\varrho_n$ . By orthogonality,

$$\begin{aligned} \Phi_n(z, w_1) &= \varrho_n \phi_n(z, w_2) + \sum_{j=0}^{n-1} \phi_j(z, w_2) \int_{-\pi}^{\pi} \Phi_n(e^{i\theta}, w_1) \bar{\phi}_j(e^{i\theta}, w_2) (w_2(\theta) - w_1(\theta)) d\theta = \\ &= \varrho_n \phi_n(z, w_2) + \int_{-\pi}^{\pi} \Phi_n(e^{i\theta}, w_1) K_{n-1}(e^{i\theta}, z, w_2) (w_2(\theta) - w_1(\theta)) d\theta. \end{aligned}$$

The Christoffel-Darboux kernel  $K_{n-1}$  admits a representation (see, e.g., [15], p. 225, formula (8.2.1)):

$$K_{n-1}(\xi, z, \mu) = \frac{\phi_n^*(z, \mu) \bar{\phi}_n^*(\xi, \mu) - \phi_n(z, \mu) \bar{\phi}_n(\xi, \mu)}{1 - z \bar{\xi}}.$$

Then,

$$|\Phi_n(1, w_1)| \leq |\phi_n(1, w_2)| \left( \varrho_n + 4\epsilon^{-1} \int_{|\theta| > \epsilon} |\Phi_n(e^{i\theta}, w_1)| \cdot |\phi_n(e^{i\theta}, w_2)| (w_1 + w_2) d\theta \right).$$

We will now use (5.1). Comparing the coefficients in front of  $z^n$  in (5.4), we get

$$\lambda(w_2) \leq \varrho_n \leq \Lambda(w_2)$$

and so

$$\left| \frac{\phi_n(1, w_1)}{\phi_n(1, w_2)} \right| \leq \frac{\Lambda(w_2)}{\lambda(w_1)} + \frac{4\Lambda(w_1)}{\epsilon \lambda(w_1)} \left( \int_{|\theta| > \epsilon} |\phi_n(e^{i\theta}, w_1) \phi_n(e^{i\theta}, w_2)| (w_1 + w_2) d\theta \right)$$

by the repetitive application of (5.1). □

## 6. THE PROOFS OF THEOREM 0.3 AND THEOREM 0.4.

We start with a Lemma which will immediately imply Theorem 0.3. It allows to perturb very general measures and have the orthogonal polynomial grow.

**Lemma 6.1.** *Assume  $\delta \in (0, 1]$ ,  $p \in [2, \infty)$  and the weight  $w$  satisfies the following properties:*

$$\|w\|_1 = 1, \quad w \geq \delta/(2\pi), \quad w \in L^p[-\pi, \pi].$$

*Then, for arbitrary  $\epsilon > 0$  and  $n \in \mathbb{N}$ , there is a weight  $\tilde{w}$  such that*

$$\|\tilde{w}\|_1 = 1, \quad \tilde{w} \geq \delta/(2\pi) - \epsilon, \quad \|w - \tilde{w}\|_p \leq \epsilon \quad (6.1)$$

*and*

$$|\phi_n(1, \tilde{w})| > C(\epsilon, \delta, p, w) \sqrt{n}.$$

*Proof.* Take any  $\delta \in (0, 1]$ . For every  $p_1 \in (p, \infty)$ , Lemma 4.1 yields  $\sigma_1 : d\sigma_1 = \sigma'_1 d\theta$  so that

$$\sigma'_1 \geq \delta_0/(2\pi) > 0, \quad \|\sigma'_1\|_{L^{p_1}(\mathbb{T})} < C(p_1), \quad \|\sigma'_1\|_{L^1(\mathbb{T})} = 1, \quad |\phi_n(1, \sigma_1)| > C(p_1) \sqrt{n}.$$

The constants  $C(p_1)$  above are  $n$ -independent. Consider an interval  $I_\tau = (-\tau, \tau)$ . Then,

$$\int_{I_\tau} d\sigma_1 \lesssim \tau^{1/p'_1}.$$

We now introduce two new weights  $w_2, \tilde{w}$  given by:

$$w_2 = \begin{cases} \sigma'_1/\delta_0, & \theta \in I_\tau \\ w, & \theta \notin I_\tau \end{cases}, \quad \tilde{w} = w_2/\|w_2\|_1.$$

We have  $w_2 \geq (2\pi)^{-1}\delta$  a.e. on  $\mathbb{T}$  and

$$\|w - w_2\|_p \leq \|w\|_{L^p[-\tau, \tau]} + \|w_2\|_{L^p[-\tau, \tau]} \leq o(1) + C(p_1)\tau^{p^{-1}-p_1^{-1}}$$

by Hölder's inequality. Here  $o(1) \rightarrow 0$  as  $\tau \rightarrow 0$ . Therefore,

$$\|w - w_2\|_1 \lesssim o(1) + C(p_1)\tau^{p^{-1}-p_1^{-1}}.$$

The triangle inequality and normalization  $\|w\|_1 = 1$  give

$$\|w_2\|_1 = 1 + o(1) + O(\tau^{p^{-1}-p_1^{-1}}).$$

We can choose  $\tau$  small enough that the last two conditions in (6.1) are satisfied for  $\tilde{w}$ . For the corresponding polynomials, we have

$$\phi_n(1, \sigma_1/\delta_0) = \sqrt{\delta_0}\phi_n(1, \sigma_1)$$

so  $|\phi_n(1, \sigma_1/\delta_0)| \gtrsim \sqrt{n}$ . Apply Lemma 5.1 with  $w_1 = \sigma'_1/\delta_0$  and  $w_2$  defined above. Notice that  $0 < C_1(\delta, p) \leq \lambda(w_{1(2)}) \leq C_2(\delta, p)$ ,  $C_1(\delta, p) \leq \Lambda(w_{1(2)}) \leq C_2(\delta, p)$  and  $w_1 = w_2$  on  $I_\tau$  by construction. We have

$$\int_{|\theta|>\tau} |\phi_n(e^{i\theta}, w_1)\phi_n(e^{i\theta}, w_2)|w_2 d\theta \lesssim \max_{I_\tau^c} |\phi_n(e^{i\theta}, w_1)| \cdot \|w_2\|_{L^2(\mathbb{T})} \cdot \|\phi_n(e^{i\theta}, w_2)\|_{L^2(\mathbb{T})}$$

and

$$\int_{|\theta|>\tau} |\phi_n(e^{i\theta}, w_1)\phi_n(e^{i\theta}, w_2)|w_1 d\theta \leq \max_{I_\tau^c} |\phi_n(e^{i\theta}, w_1)| \cdot \|w_1\|_{L^2(\mathbb{T})} \cdot \|\phi_n(e^{i\theta}, w_2)\|_{L^2(\mathbb{T})}.$$

The polynomials orthonormal with respect to a measure in Steklov class have uniformly bounded  $L^2(\mathbb{T})$  norm (see the proof of Lemma 0.1). For the estimation of  $\max_{I_\tau^c} |\phi_n(e^{i\theta}, w_1)|$ , we use (3.49) to get

$$\max_{I_\tau^c} |\phi_n(e^{i\theta}, w_1)| \lesssim \left( \frac{n}{1+n^2\tau^2} + \tau^{-\alpha} \right)^{1/2} \lesssim \tau^{-1/2}$$

where  $\alpha < 1$ . Thus, the localization principle (5.3) gives

$$|\phi_n(1, \tilde{w})| \geq C(\epsilon, \delta, p, w)\sqrt{n}$$

and the proof is finished.  $\square$

Now the proof of the Theorem 0.3 is immediate.

*Proof. (of Theorem 0.3)* It is sufficient to take  $w = (2\pi)^{-1}$  and  $\delta = 1$ .  $\square$

**Remark.** Notice that this proof allows to improve Lemma 4.1 to cover the full range of  $\delta : \delta \in (0, 1)$ . This statement is much stronger than the Theorem 0.3 itself: it shows that  $\sqrt{n}$  growth can be achieved on far more regular weights.

Now, we can iterate this construction to prove Theorem 0.4.

*Proof. (of Theorem 0.4).* Fix any  $\delta \in (0, 1)$  and a sequence  $\{\beta_n\} : \lim_{n \rightarrow \infty} \beta_n = 0$ . We can assume without loss of generality that  $\beta_1 < 1$ . Choose any  $p \in [2, \infty)$  and parameter  $\tilde{C} > 1$ . We construct the sequence of weights  $\{w_n\}$  through the following induction:

- **First step:** We let  $w_1 = (2\pi)^{-1}$  and  $k_1 = 1$ . Then,  $|\phi_{k_1}(1, w_1)| = 1 > \beta_1$ .

- **Inductive assumption:** We assume that the weight  $w_n$  and the natural numbers  $k_1 < \dots < k_n$  are given so that

$$|\phi_{k_j}(1, w_n)| > \beta_{k_j} \sqrt{k_j}, \quad j = 1, \dots, n$$

and

$$\|w_n\|_1 = 1, \quad \|w_n\|_p < \tilde{C}, \quad w_n \geq (\delta + (1 - \delta)2^{-n})/(2\pi). \quad (6.2)$$

- **Inductive step:** For every  $\epsilon > 0$  and  $N$  we can use the perturbation Lemma 6.1 to get  $w_{n+1}$  so that

$$\|w_{n+1}\|_1 = 1, \quad w_{n+1} \geq (\delta + (1 - \delta)2^{-n})/(2\pi) - \epsilon, \quad \|w_{n+1} - w_n\|_p \leq \epsilon$$

and

$$|\phi_N(1, w_{n+1})| > C(\epsilon, \delta, p) \sqrt{N}.$$

Notice that for fixed  $j$  the functional  $\phi_j(1, \sigma)$  is continuous in  $\sigma$  in weak- $(*)$  (and then in  $L^p(\mathbb{T})$ ) topology. The second inequality in (6.2) is strict. So, we first choose  $\epsilon$  so small that:

1.  $\|w_n\|_p + \epsilon < \tilde{C}$ .
2.  $|\phi_{k_j}(1, \nu)| > \beta_{k_j} \sqrt{k_j}$ ,  $\forall j = 1 \dots n$ , as long as  $\nu : \|\nu - w_n\|_p < \epsilon$ .
3.  $(\delta + (1 - \delta)2^{-n})/(2\pi) - \epsilon > (\delta + (1 - \delta)2^{-n-1})/(2\pi)$ .

Then, with fixed  $\epsilon$ , take  $N$  large so that  $N > k_n$  and  $C(\epsilon, \delta, p) \sqrt{N} > \beta_N \sqrt{N}$ . We can always achieve that since  $\lim_{l \rightarrow \infty} \beta_l = 0$ . Now, let  $k_{n+1} = N$ .

Thus, we constructed the new weight  $w_{n+1}$  and  $k_{n+1}$  that satisfy all induction assumptions. At each step when going from  $n$  to  $n + 1$  we choose new  $\epsilon$  that depends on  $n$ , the step of induction. By construction,  $\|w_{n+1} - w_n\|_p \lesssim 2^{-n}$  so  $w_n$  converges to some  $w$  in  $L^p(\mathbb{T})$  norm. Moreover,  $w \geq \delta/(2\pi)$  a.e. on  $\mathbb{T}$ . We use the continuity of  $\phi_j(1, \sigma)$  in  $\sigma$  again to get

$$|\phi_{k_j}(1, w)| \geq \beta_{k_j} \sqrt{k_j}, \quad \forall j$$

and that finishes the proof.  $\square$

**Remark.** It is clear that our construction allows to have the polynomials grow simultaneously at any finite number of points on the circle. We also can make the measure of orthogonality symmetric with respect to both axis  $OX$  and  $OY$ . Indeed, the measure we constructed in the Theorem is given by the even weight  $w$ . Now, for every  $N \in \mathbb{N}$ , we can take  $w_N(x) = w(Nx)$  and then

$$\phi_{Nj}(z, w_N) = \phi_j(z^N, w)$$

To make the measure symmetric with respect to both  $OX$  and  $OY$ , it is sufficient to take  $N = 2$ .

**Remark.** The conjecture of Steklov and its solution can be interpreted as follows. It is known that  $\{\Phi_n(z)\}$  satisfy the recursion

$$\Phi_{n+1}^*(z) = \Phi_n^*(z) - \gamma_n z \Phi_n(z), \quad \Phi_0^*(z) = 1.$$

Therefore,

$$\Phi_n^*(z) = 1 - z \sum_{j=0}^{n-1} \gamma_j \Phi_j(z).$$

Recall that  $\sigma \in S_\delta$  implies  $\{\gamma_j\} \in \ell^2$  and one can define a maximal function in analogy to the Carleson maximal function in Fourier series, i.e.,

$$\mathcal{M}(\theta) = \sup_n \left| \sum_{j=0}^n \gamma_j \Phi_j(e^{i\theta}) \right|, \quad \mathcal{M}(\theta) \sim \sup_n |\Phi_n^*(e^{i\theta}) - 1|.$$

Then, for the example we constructed,

$$\sum_{j=0}^{\infty} \gamma_j \Phi_j(e^{i\theta}) \in L^\infty(\mathbb{T}) \quad \text{but} \quad \mathcal{M}(e^{i\theta}) \notin L^\infty(\mathbb{T}).$$

## Part 4. Applications

In this part, we apply the obtained results to handle the case of the orthogonality on the segment on the real line. We also prove the sharp bounds for the polynomial entropy in the Steklov class.

### 7. BACK TO THE REAL LINE.

In the case when the measure  $\sigma$  is symmetric on  $\mathbb{T}$  with respect to the real line, one can relate  $\{\phi_n(z, \sigma)\}$  to polynomials orthogonal on the real line through the following standard procedure. Let  $\psi, (x \in [-1, 1], \psi(-1) = 0)$  be a non-decreasing bounded function with an infinite number of growth points. Consider the system of polynomials  $\{P_k\}, (k = 0, 1, \dots)$  orthonormal with respect to the measure  $\psi$  supported on the segment  $[-1, 1]$ . Introduce the function

$$\sigma(\theta) = \begin{cases} -\psi(\cos \theta), & 0 \leq \theta \leq \pi, \\ \psi(\cos \theta), & \pi \leq \theta \leq 2\pi, \end{cases} \quad (7.1)$$

which is bounded and non-decreasing on  $[0, 2\pi]$ . Consider the polynomials  $\phi_k(z, \sigma) = \lambda_k z^k + \dots$ ,  $(\lambda_k = (\rho_0 \dots \rho_{k-1})^{-1} > 0)$  orthonormal with respect to  $\sigma$ . Then,  $\phi_n$  is related to  $P_k$  by the formula

$$P_k(x, \psi) = \frac{\phi_{2k}(z, \sigma) + \phi_{2k}^*(z, \sigma)}{\sqrt{2\pi [1 + \lambda_{2k}^{-1} \phi_{2k}(0, \sigma)]}} z^{-k}, \quad k = 0, 1, \dots, \quad (7.2)$$

where  $x = (z + z^{-1})/2$  ([10, 24]). This reduction also works in the opposite direction: given the symmetric measure  $\sigma$  we can map it to the measure on the real line and the corresponding polynomials will be related by (7.2). We are ready to formulate the Theorem.

**Theorem 7.1.** *Let  $\delta > 0$ . Then, for every positive sequence  $\{\beta_n\} : \lim_{n \rightarrow \infty} \beta_n = 0$ , there is a measure  $\psi : d\psi = \psi' dx$  supported on  $[-1, 1]$  such that  $\psi'(x) \geq \delta$  for a.e.  $x \in [-1, 1]$  and*

$$|P_{k_n}(0, \psi)| \gtrsim \beta_{k_n} \sqrt{k_n} \quad (7.3)$$

for some sequence  $\{k_n\} \subset \mathbb{N}$ .

*Proof.* Indeed, in the Theorem 0.4 we can take  $\sigma^* : d\sigma^* = w^* d\theta$  to be symmetric with respect to both axis, i.e.,  $w^*(2\pi - \theta) = w^*(\theta)$  (symmetry with respect to  $OX$ ) and  $w^*(\theta) = w^*(\pi - \theta)$  (symmetry with respect to  $OY$ ). Moreover, we can always arrange for all  $\{k_n\}$  to be divisible by 4 and

$$|\phi_{k_n}(1, \sigma^*)| \geq \beta_{k_n} \sqrt{k_n}$$

Now, we take  $\sigma : d\sigma = w^*(\theta - \pi/2) d\theta$ , i.e., the rotation of  $\sigma^*$  by  $\pi/2$  and apply (7.1) to it. The symmetries of  $\sigma^*$  yield the symmetry of  $\sigma$  with respect to  $OX$  so this transform is applicable. Notice that  $\phi_n(z, \sigma) = e^{in\pi/2} \phi_n(z e^{-i\pi/2}, \sigma^*)$  where the first factor is introduced to make the leading coefficient positive. Also, notice that  $\phi_n^*(1, \sigma^*)$  is real-valued so  $\phi_n(1, \sigma^*) = \phi_n^*(1, \sigma^*)$ . We have

$$\begin{aligned} \phi_{k_n}(i, \sigma) + \phi_{k_n}^*(i, \sigma) &= \phi_{k_n}(i, \sigma) + \overline{\phi_{k_n}(i, \sigma)} = \\ \phi_{k_n}(1, \sigma^*) + \overline{\phi_{k_n}(1, \sigma^*)} &= 2\phi_{k_n}(1, \sigma^*) \geq 2\beta_{k_n} \sqrt{k_n} \end{aligned}$$

Now, notice that

$$\phi_{2k}(0, \sigma)/\lambda_{2k} = \Phi_{2k}(0, \sigma) = -\overline{\gamma}_{2k-1} .$$

By the Szegő sum rule,  $\gamma_n \in \ell^2$  and so  $\lim_{n \rightarrow \infty} \gamma_n = 0$ . Thus, (7.2) gives

$$|P_{k_n/2}(0, \psi)| \gtrsim \beta_{k_n} \sqrt{k_n}$$

and, after redefining  $k_n$ , this is exactly (7.3). For the derivative of  $\psi$ , we have

$$\psi'(\cos \theta) = \frac{\sigma'(\theta)}{|\sin \theta|} > \frac{\delta_1/(2\pi)}{|\sin \theta|} > \delta_1/(2\pi) = \delta$$

if  $\delta_1 = 2\pi\delta$ . □

**Remark.** The original conjecture of Steklov was formulated in terms of the weights (i.e., the unit ball in  $L^1(\mathbb{T})$ ) and we solved it in that form. However, as the results on maximizers from the first part of the paper suggest, the class of probability measures is far more natural for that setting.

## 8. THE POLYNOMIAL ENTROPIES AND THE STEKLOV CLASS.

In recent years, a lot of efforts were made (see, e.g., [2, 3, 5]) to study the so-called polynomial entropy

$$\int_{\mathbb{T}} |\phi_n|^2 \ln |\phi_n| d\sigma$$

where  $\phi_n$  are orthonormal with respect to  $\sigma$ . Since  $\sup_{x \in [0,1]} x^2 |\ln x| < \infty$ , this quantity is bounded if and only if

$$\int_{\mathbb{T}} |\phi_n|^2 \ln^+ |\phi_n| d\sigma$$

is bounded. The last expression is important as it contains the information on the size of  $\phi_n$ . In this section, we consider the following variational problem

$$\Omega_n(\mathcal{K}) = \sup_{\sigma \in \mathcal{K}} \int_{\mathbb{T}} |\phi_n|^2 \ln^+ |\phi_n| d\sigma$$

where  $\phi_n$  is the  $n$ -th orthonormal polynomial with respect to  $\sigma$  taken in  $\mathcal{K}$ , some special class of measures. It is an interesting question to describe those  $\mathcal{K}$  for which  $\Omega_n(\mathcal{K})$  is bounded in  $n$ . So far, this is known only for very few  $\mathcal{K}$ , e.g., the Baxter class of measures. For the Szegő class with measures normalized by the  $\ell^2$  norm of Schur parameters, the sharp estimate  $\Omega_n \sim \sqrt{n}$  is known [8]. In this section, we will obtain the sharp bound on  $\Omega_n(S_\delta)$ .

**Lemma 8.1.** *If  $\delta \in (0, 1)$ , then*

$$\Omega_n(S_\delta) \sim \ln n .$$

*Proof.* If one takes the measure  $\sigma$  and the polynomial  $\phi_n$  constructed in the proof of the Theorem 0.3, then

$$\Omega_n(S_\delta) \gtrsim \int_{\mathbb{T}} |\phi_n|^2 \ln^+ |\phi_n| d\theta \gtrsim 1 + \int_{0.01n^{-1}}^{\theta_1 - 0.01n^{-1}} |f_n|^2 \ln^+ |f_n| d\theta$$

where  $\theta_1$  was introduced in this proof. On that interval,  $|\Psi(\theta) + \sin(n\theta - 2\phi)| > C$  and so  $|f_n| \sim |Q_m|$ . This follows from (3.30) and the verification of the normalization condition. Then, the expression (8.14) gives a very rough lower bound

$$|Q_m(e^{i\theta})|^2 \gtrsim \frac{m}{m^2\theta^2 + 1} + 1 .$$

This shows  $|Q_m| \sim \sqrt{n}$  on the interval  $(0.01n^{-1}, \theta_1 - 0.01n^{-1})$  and so

$$\Omega_n(S_\delta) \gtrsim n^{-1} \cdot n \cdot \ln n \sim \ln n .$$

Therefore, the polynomial entropy grows at least as the logarithm. On the other hand, the trivial upper bound  $\|\phi_n\|_\infty \lesssim \sqrt{n}$  implies that  $\Omega_n(S_\delta) \lesssim \ln n$ .  $\square$

## Some open problems

In conclusion, we want to discuss some interesting problems we didn't address.

- (1) In the variational problem for  $M_{n,\delta}$ , it would be interesting to know whether the maximizer is unique and how many mass points it possesses. Ideally, one would want to find it explicitly. At the moment, very little is known about the maximizers  $\mu^*$  in Theorem 1.2. In [7], it was proved that  $N$ , the number of point masses, is of order  $n$ .
- (2) Suppose that  $\phi_n(z, \sigma)$  is the orthonormal polynomial,  $\sigma \in S_\delta$ , and  $|\phi_n(1, \sigma)| > C\sqrt{n}$ . What is the behavior of the Schur parameters  $\{\gamma_j(\sigma)\}$ ? This question is interesting as its answer can give a “difference equation perspective” to the problem. To this end, one only needs to find the coefficients of the Szegő recursion (Schur parameters) such that

$$\limsup_{j \rightarrow \infty} \|\phi_j(e^{i\theta}, \sigma)\|_{L^\infty(\mathbb{T})} < C$$

(which is equivalent to the Steklov condition if  $\sigma$  is regular) but

$$\|\phi_n(e^{i\theta}, \sigma)\|_{L^\infty(\mathbb{T})} \sim \sqrt{n} .$$

for the fixed arbitrarily large  $n$ .

- (3) For the following variational problem

$$M_{n,\delta_1,\delta_2} = \sup_{\delta_1/(2\pi) \leq \sigma' \leq \delta_2/(2\pi)} \|\phi_n(e^{i\theta}, \sigma)\|_{L^\infty(\mathbb{T})}$$

find the sharp estimates for  $M_{n,\delta_1,\delta_2}$  as  $n \rightarrow \infty$ .

## Acknowledgement.

The research of S.D. was supported by NSF grant DMS-1067413. The research of A.A. and D.T. was supported by the grants RFBR 13-01-12430 OFIm and RFBR 11-01-00245 and Program 1 DMS RAS. The hospitality of IMB (Bordeaux) and IHES (Paris) is gratefully acknowledged by S.D. The authors thank Stas Kupin, Fedor Nazarov, and Evgenii Rakhmanov for interesting comments.

## Appendix A.

In this Appendix, we start by introducing the polynomials that approximate the function  $(1-z)^{-\alpha/2}$  (used in the formula (3.26)) and the function  $(1-z)^{1-\alpha}$  (used in the definition of  $P_m$ , formula (3.29)). These polynomials are well studied (see, e.g., [30], chapter 5) but we deduce the



necessary estimates here for completeness of exposition. Notice first, that  $(1 - z)^\beta$  is analytic in  $\mathbb{D}$  and has positive real part for any  $\beta \in (-1, 1)$ . For  $z = e^{i\theta} \in \mathbb{T}$ , we have

$$(1 - z)^\beta = \left( (1 - \cos \theta)^2 + \sin^2 \theta \right)^{\beta/2} \exp(-i\beta L(\theta)) = \left( 2 \sin \frac{|\theta|}{2} \right)^\beta \exp(-i\beta L(\theta))$$

where

$$L(\theta) = \arctan \left( \frac{\sin \theta}{1 - \cos \theta} \right) = \arctan \left( \cot \frac{\theta}{2} \right)$$

and so

$$(1 - z)^\beta = |\theta|^\beta (1 + O(\theta^2)) \exp(-i\beta L(\theta)), \quad \theta \rightarrow 0; \quad L(\theta) = \text{sign}(\theta) \frac{\pi - |\theta|}{2}. \quad (8.1)$$

We will now introduce the polynomials that approximate  $(1 - z)^\beta$  uniformly on compacts in  $\mathbb{D}$  and behave on the boundary in a controlled way. We will treat the cases of positive and negative  $\beta$  separately. Let  $A_n(z)$  be the  $n$ -th Taylor polynomial of  $(1 - z)^\beta$  with positive  $\beta$ , i.e.,

$$A_n(z) = 1 - \sum_{j=1}^n c_j z^j, \quad c_j = c_j(\beta) = \frac{\beta(1-\beta) \dots (j-1-\beta)}{j!}.$$

The polynomial  $R_{(n, -(1-\alpha))}$  in the main text will be taken as  $A_n$  with  $\beta = 1 - \alpha \in (0, 1/2)$ . For  $B_n(z)$ , we choose  $n$ -th Taylor coefficient of  $(1 - z)^{-\beta}$  with positive  $\beta$ , i.e.,

$$B_n(z) = 1 + \sum_{j=1}^n d_j z^j$$

and

$$d_j = d_j(\beta) = \frac{\beta(\beta+1) \dots (\beta+j-1)}{j!} = \frac{j^{\beta-1}}{\Gamma(\beta)} + O(j^{\beta-2}). \quad (8.2)$$

The polynomial  $R_{(n, \alpha/2)}$  used in the main text is  $B_n$  with  $\beta = \alpha/2 \in (1/4, 1/2)$ . We need the following simple Lemmas.

**Lemma 8.2.** *For any  $a > 0$ , we have*

$$\int_0^a \frac{\cos x}{x^\gamma} dx > 0, \quad \text{if } \gamma \in [1/2, 1)$$

and

$$\int_0^a \frac{\sin x}{x^\gamma} dx > 0, \quad \int_0^\infty \frac{\sin x}{x^\gamma} dx > 0, \quad \text{if } \gamma \in (0, 1).$$

*Proof.* The inequalities with  $\sin$  are elementary as  $x^{-\gamma}$  decays and  $\sin x$  satisfies

$$\sin(\pi + x) = -\sin x; \quad \sin x > 0, \quad x \in (0, \pi)$$

For the first inequality, we notice that

$$\int_{3\pi/2}^a \frac{\cos x}{x^\gamma} dx > 0$$

for any  $a > 3\pi/2$  and we only need to show that

$$\int_0^{3\pi/2} \frac{\cos x}{x^\gamma} dx > 0.$$

Integrating by parts we have

$$\int_0^{3\pi/2} \frac{\cos x}{x^\gamma} dx = -\left(\frac{2}{3\pi}\right)^\gamma + \gamma \int_0^{3\pi/2} \frac{\sin x}{x^{\gamma+1}} dx > -\left(\frac{2}{3\pi}\right)^\gamma + \frac{2\gamma}{\pi} \int_0^{\pi/2} x^{-\gamma} dx$$

where we dropped the integral over  $[\pi/2, 3\pi/2]$  in the last inequality and used the fact that  $x^{-1} \sin x$  decays monotonically on  $[0, \pi/2]$ . Calculating the integral, we get

$$\frac{2\gamma}{\pi(1-\gamma)} \left(\frac{\pi}{2}\right)^{1-\gamma} - \left(\frac{2}{3\pi}\right)^\gamma = \left(\frac{2}{\pi}\right)^\gamma \left(\frac{\gamma}{1-\gamma} - 3^{-\gamma}\right) > 0$$

for  $\gamma \in [1/2, 1)$ . □

Let us first study the properties of  $B_n$ . As  $B_n$  is the Taylor expansion of  $(1-z)^{-\beta}$  and  $\beta \in (0, 1/2)$ , we have the uniform convergence  $B_n(z) \rightarrow (1-z)^{-\beta}$  in  $\{|z| \leq 1\} \cap \{|1-z| > 1-v\}$  for any fixed  $v > 0$  as long as  $n \rightarrow \infty$ . Indeed, due to monotonicity of  $d_j$  we have

$$\sum_{j=1}^n |d_{j+1} - d_j| = |d_{n+1} - d_1|.$$

Then, the Abel's transform yields the uniform convergence.

We now take  $z = e^{i\theta}$  with  $\theta \in (-v, v)$  where  $v$  is small.

We will need to use the following approximations by the integrals. Let  $\gamma \in (0, 1)$ .

$$\int_1^n \frac{\cos(x\theta)}{x^\gamma} dx = \sum_{j=1}^{n-1} \int_j^{j+1} \frac{\cos(x\theta)}{x^\gamma} dx = \sum_{j=1}^{n-1} \frac{1}{j^\gamma} \int_j^{j+1} \cos(x\theta) dx + \sum_{j=1}^{n-1} \int_j^{j+1} \cos(x\theta) \left(\frac{1}{x^\gamma} - \frac{1}{j^\gamma}\right) dx. \quad (8.3)$$

Since

$$\max_{x \in [j, j+1]} |x^{-\gamma} - j^{-\gamma}| \lesssim j^{-\gamma-1} \quad (8.4)$$

the second term is  $O(1)$  uniformly in  $\theta$  and  $n$  and that gives

$$\begin{aligned} \int_1^n \frac{\cos(x\theta)}{x^\gamma} dx &= O(1) + \sum_{j=1}^{n-1} \frac{1}{j^\gamma} \frac{\sin(\theta/2)}{\theta/2} \cos(j\theta + \theta/2) = \\ &= O(1) + \sum_{j=1}^{n-1} \frac{1}{j^\gamma} \frac{\sin(\theta/2)}{\theta/2} \left( \cos(j\theta) \cos(\theta/2) - \sin(j\theta) \sin(\theta/2) \right). \end{aligned}$$

Similarly

$$\int_1^n \frac{\sin(x\theta)}{x^\gamma} dx = \sum_{j=1}^{n-1} \int_j^{j+1} \frac{\sin(x\theta)}{x^\gamma} dx = \sum_{j=1}^{n-1} \frac{1}{j^\gamma} \int_j^{j+1} \sin(x\theta) dx + \sum_{j=1}^{n-1} \int_j^{j+1} \sin(x\theta) \left(\frac{1}{x^\gamma} - \frac{1}{j^\gamma}\right) dx. \quad (8.5)$$

By (8.4), the second term is  $o(1)$  as  $\theta \rightarrow 0$ , uniformly in  $n$ . Therefore, we have

$$\begin{aligned} \int_1^n \frac{\sin(x\theta)}{x^\gamma} dx &= o(1) + \sum_{j=1}^{n-1} \frac{1}{j^\gamma} \int_j^{j+1} \sin(x\theta) dx = \\ &= o(1) + \sum_{j=1}^{n-1} \frac{1}{j^\gamma} \frac{\sin(\theta/2)}{\theta/2} \sin(j\theta + \theta/2) = \\ &= o(1) + \sum_{j=1}^{n-1} \frac{1}{j^\gamma} \frac{\sin(\theta/2)}{\theta/2} \left( \sin(j\theta) \cos(\theta/2) + \cos(j\theta) \sin(\theta/2) \right). \end{aligned}$$

Above,  $O(1)$  and  $o(1)$  are written for  $\theta \rightarrow 0$  and they are uniform in  $n$ . Now, representations (8.3) and (8.5) yield the formulas for

$$\sum_{j=1}^{n-1} \frac{\cos(j\theta)}{j^\gamma}, \quad \sum_{j=1}^{n-1} \frac{\sin(j\theta)}{j^\gamma}$$

i.e.,

$$\sum_{j=1}^{n-1} \frac{\cos(j\theta)}{j^\gamma} = O(1) + C_{11}(\theta) \int_1^n \frac{\cos(x\theta)}{x^\gamma} dx + C_{12}(\theta) \int_1^n \frac{\sin(x\theta)}{x^\gamma} dx \quad (8.6)$$

and

$$\sum_{j=1}^{n-1} \frac{\sin(j\theta)}{j^\gamma} = o(1) + C_{21}(\theta) \int_1^n \frac{\sin(x\theta)}{x^\gamma} dx + C_{22}(\theta) \int_1^n \frac{\cos(x\theta)}{x^\gamma} dx \quad (8.7)$$

where  $C_{11} \rightarrow 1$ ,  $C_{12} \rightarrow 0$ ,  $C_{21} \rightarrow 1$ ,  $C_{22} \rightarrow 0$  as  $\theta \rightarrow 0$  uniformly in  $n$ .

Now we are ready for the next Lemma.

**Lemma 8.3.** *Let  $\beta \in (0, 1/2)$  and  $v$  is sufficiently small fixed positive number, then*

$$\operatorname{Re} B_n(e^{i\theta}) \sim (n^{-1} + |\theta|)^{-\beta}, \quad \theta \in (-v, v)$$

and

$$\begin{aligned} \frac{\operatorname{Im} B_n(e^{i\theta})}{\operatorname{sign}(\theta)} &\sim |\theta|^{-\beta}, \quad 0.01n^{-1} < |\theta| < v, \\ \frac{\operatorname{Im} B_n(e^{i\theta})}{\theta} &\sim n^{1+\beta}, \quad |\theta| < 0.01n^{-1}. \end{aligned}$$

*Proof.* The case  $|\theta| < 0.01n^{-1}$  follows from (8.2) since  $\cos(j\theta) \sim 1$  and  $\sin(j\theta)/(j\theta) \sim 1$ . For the other  $\theta$ , we first notice that it is sufficient to consider  $\theta \in (0.01n^{-1}, v)$  and that (8.2) gives

$$B_n(e^{i\theta}) = 1 + \frac{1}{\Gamma(\beta)} \left( \sum_{j=1}^n j^{-1+\beta} e^{i\theta j} + O(1) \right).$$

Let  $\gamma = 1 - \beta \in (1/2, 1)$  and use the formulas (8.6) and (8.7). Notice that

$$\int_1^n \frac{\cos(x\theta)}{x^\gamma} dx = \theta^{\gamma-1} \int_\theta^{n\theta} \frac{\cos t}{t^\gamma} dt \sim \theta^{\gamma-1} \quad (\text{any } \gamma \in (1/2, 1))$$

as long as  $\theta \in (0.01n^{-1}, v)$ . That follows from the Lemma 8.2. The last estimate is valid for sufficiently small  $v$ . Indeed, taking  $\mathcal{F}$  as

$$\mathcal{F}(t) = \int_0^t \frac{\cos u}{u^\gamma} du,$$

we get the following bounds

$$\int_\theta^{n\theta} \frac{\cos t}{t^\gamma} dt = \mathcal{F}(n\theta) - \mathcal{F}(\theta) \geq \min_{[0.01, \infty]} \mathcal{F} - \max_{[0, v]} \mathcal{F} \geq \text{const} > 0.$$

Similarly

$$\int_1^n \frac{\sin(x\theta)}{x^\gamma} dx = \theta^{\gamma-1} \int_\theta^{n\theta} \frac{\sin t}{t^\gamma} dt \sim \theta^{\gamma-1} \quad (\text{any } \gamma \in (0, 1)).$$

That finishes the proof. □

**Lemma 8.4.** For any  $\beta \in (0, 1)$ , we have

$$|B'_n(e^{i\theta})| \lesssim \begin{cases} |\theta|^{-1} n^\beta, & |\theta| > n^{-1} \\ n^{1+\beta}, & |\theta| < n^{-1} \end{cases}$$

$$|B''_n(e^{i\theta})| \lesssim \begin{cases} |\theta|^{-1} n^{\beta+1}, & |\theta| > n^{-1} \\ n^{2+\beta}, & |\theta| < n^{-1} \end{cases}$$

where the derivative is taken in  $\theta \in (-v, v)$ .

*Proof.* For  $|\theta| < n^{-1}$ , this follows from

$$B'_n = \sum_{j=1}^n i j d_j e^{ij\theta}, \quad B''_n = \sum_{j=1}^n (ij)^2 d_j e^{ij\theta}$$

by estimating the absolute values of each term.

For  $|\theta| > n^{-1}$ , we can use Abel's Lemma. Indeed,

$$|B'_n| = \left| \sum_{j=1}^n j d_j e^{ij\theta} \right| \lesssim \left| \sum_{j=1}^n e^{ij\theta} j^\beta \right| + \sum_{j=1}^n j^{\beta-1}.$$

The second term in the sum is bounded by  $Cn^\beta$ . For the first one, we have

$$\left| \sum_{j=1}^n e^{ij\theta} j^\beta \right| \lesssim n^\beta |S_n| + \left| \sum_{j=1}^n S_j j^{\beta-1} \right|, \quad S_j = \sum_{k=1}^j e^{ik\theta}, \quad |S_j| \lesssim |\theta|^{-1}$$

and that yields the bound for  $B'_n$ . The second derivative can be estimated similarly.  $\square$

Next, we will study the polynomial  $A_n$ . For the Taylor expansion of  $(1-z)^\beta$ , we have

$$(1-z)^\beta = 1 + \sum_{j=1}^{\infty} \frac{(-1)^j \beta(\beta-1) \dots (\beta-(j-1))}{j!} z^j = 1 - \sum_{j=1}^{\infty} c_j z^j.$$

The coefficients behave as follows

$$c_j = \frac{\beta(1-\beta) \dots (j-1-\beta)}{j!} = \frac{-j^{-\beta-1}}{\Gamma(-\beta)} + O(j^{-\beta-2}) > 0. \quad (8.8)$$

The series  $\sum_j c_j$  converges absolutely and

$$1 - \sum_{j=1}^n c_j z^j \rightarrow (1-z)^\beta, \quad |z| < 1$$

therefore

$$1 - \sum_{j=1}^{\infty} c_j z^j = (1-z)^\beta$$

on  $\mathbb{T}$  pointwise. In particular,

$$\sum_{j=1}^{\infty} c_j = 1.$$

Therefore, the formula for  $A_n$  can be rewritten as

$$A_n(z) = M_n + \sum_{j=1}^n c_j (1-z^j), \quad M_n = \sum_{j=n+1}^{\infty} c_j = \frac{n^{-\beta}}{\Gamma(1-\beta)} + O(n^{-\beta-1}).$$

We again notice that  $A_n(z)$  converges to  $(1-z)^\beta$  uniformly in  $\{|z| \leq 1\}$ . Indeed  $|A_n(z)| < 2$  uniformly in  $\mathbb{D}$  and  $|(1-z)^\beta - A_n(z)| = \left| \sum_{j=n+1}^{\infty} c_j z^j \right| < M_n$ .

**Lemma 8.5.** *Let  $\beta \in (0, 1)$ . We have*

$$\operatorname{Re} A_n(e^{i\theta}) \sim (n^{-1} + |\theta|)^\beta, \quad \theta \in (-v, v) \quad (8.9)$$

and

$$-\frac{\operatorname{Im} A_n(e^{i\theta})}{\operatorname{sign}(\theta)} \sim \begin{cases} |\theta| n^{1-\beta}, & |\theta| < 0.01 n^{-1} \\ |\theta|^\beta, & 0.01 n^{-1} < |\theta| < v \end{cases} \quad (8.10)$$

*Proof.* We only need to handle positive  $\theta$ . Again, if  $0 < \theta < 0.01 n^{-1}$ , the estimate is simple.

$$\operatorname{Re} A_n = M_n + \sum_{j=1}^n c_j (1 - \cos(j\theta))$$

and we have a bound

$$n^{-\beta} \sim M_n \leq \operatorname{Re} A_n \lesssim M_n + \sum_{j=1}^n j^{-\beta-1} (j^2 \theta^2) \lesssim n^{-\beta} + n^{-\beta}.$$

Similarly

$$\operatorname{Im} A_n = - \sum_{j=1}^n c_j \sin(j\theta)$$

and

$$\sum_{j=1}^n c_j \sin(j\theta) \sim \theta \sum_{j=1}^n j c_j \lesssim \theta n^{1-\beta}.$$

For  $\theta$  from  $[0.01 n^{-1}, v]$ , we can again approximate by the integrals. We have

$$\sum_{j=1}^n c_j \sin(j\theta) = \frac{-1}{\Gamma(-\beta)} \sum_{j=1}^n j^{-\beta-1} \sin(j\theta) + O\left(\sum_{j=1}^n j^{-\beta-2} (j\theta)\right).$$

The last term is  $O(\theta)$ . Then, take

$$\int_1^n \frac{\sin(x\theta)}{x^{1+\beta}} dx = \sum_{j=1}^{n-1} \int_j^{j+1} \frac{\sin(x\theta)}{x^{1+\beta}} dx = \sum_{j=1}^{n-1} j^{-\beta-1} \int_j^{j+1} \sin(x\theta) dx + O\left(\sum_{j=1}^{n-1} \left(\frac{1}{j^{\beta+1}} - \frac{1}{(j+1)^{\beta+1}}\right) (j+1)\theta\right).$$

The second term is

$$O\left(\sum_{j=1}^{n-1} j^{-\beta-1} \theta\right) = O(\theta).$$

For the first sum, we have

$$\begin{aligned} \sum_{j=1}^{n-1} j^{-\beta-1} \int_j^{j+1} \sin(x\theta) dx &= \sum_{j=1}^{n-1} j^{-\beta-1} \frac{\sin(\theta/2)}{\theta/2} \sin(j\theta + \theta/2) = \\ &= \frac{\sin(\theta/2) \cos(\theta/2)}{\theta/2} \sum_{j=1}^{n-1} j^{-\beta-1} \sin(j\theta) + \frac{\sin^2(\theta/2)}{\theta/2} \sum_{j=1}^{n-1} j^{-\beta-1} \cos(j\theta). \end{aligned}$$

The second term is  $O(\theta)$  and

$$\frac{\sin(\theta/2) \cos(\theta/2)}{\theta/2} \sim 1$$

for  $\theta \in (0, v)$ . Then,

$$\int_1^n \frac{\sin(x\theta)}{x^{1+\beta}} dx = \theta^\beta \int_\theta^{n\theta} \frac{\sin x}{x^{1+\beta}} dx = \int_0^{n\theta} \frac{\sin x}{x^{1+\beta}} dx + O(\theta) .$$

Notice that

$$C > \int_0^a \frac{\sin x}{x^{1+\beta}} dx > \delta_2 > 0$$

for any  $a > 0.01$  and so we have

$$\sum_{j=1}^n c_j \sin(j\theta) \sim \theta^\beta + O(\theta) \sim \theta^\beta .$$

This implies (8.10). For the real part,

$$\operatorname{Re} A_n(e^{i\theta}) = M_n + T_n(\theta) + O\left(\theta \sum_{j=1}^n j^{-1-\beta}\right), \quad T_n = \sum_{j=1}^n \frac{1 - \cos(j\theta)}{j^{1+\beta}} .$$

The last term is  $O(\theta)$ . For  $T_n$ , we have

$$T_n(0) = 0, \quad T'_n(\theta) = \sum_{j=1}^n \frac{\sin(j\theta)}{j^\beta} .$$

If  $\theta \in (0, 0.01n^{-1})$ , then  $T'_n \sim \theta n^{2-\beta}$ . For  $\theta \in (0.01n^{-1}, v)$ , the formula (8.7) gives

$$T'_n \sim \theta^{\beta-1} .$$

Integration yields

$$T_n(\theta) = \int_0^\theta T'_n(\xi) d\xi \sim \theta^\beta, \quad \theta \in (0.01n^{-1}, v) .$$

That finishes the proof. □

It is instructive to compare the results of Lemmas 8.3 and 8.5 with (8.1).

For the derivative of  $A_n$  in  $\theta$ , we have

$$A'_n = -i \sum_{j=1}^n j c_j e^{ij\theta} .$$

**Lemma 8.6.** *If  $\beta \in (0, 1)$ , then*

$$|A'_n| \lesssim \begin{cases} |\theta|^{\beta-1}, & |\theta| > 0.01n^{-1} \\ n^{1-\beta}, & |\theta| < 0.01n^{-1} \end{cases}$$

*uniformly in  $n$ .*

*Proof.* For  $|\theta| < 0.01n^{-1}$ , the estimate is obtained by taking the absolute values in the sum. For  $|\theta| > 0.01n^{-1}$ ,

$$|A'_n| \lesssim \left| \sum_{j=1}^n j^{-\beta} e^{ij\theta} \right| + 1 .$$

The estimates (8.6) and (8.7) along with the trivial bounds on the integrals involved yield the statement of the Lemma. □

**Remark.** Notice that, as  $n$  is large enough, the estimates obtained in Lemmas 8.3 and 8.5 (except for the bounds on the imaginary parts that are violated near  $z = -1$ ) can be extended from the small arc  $|\theta| < v$  to the whole circle using the uniform convergence of the corresponding Taylor expansions outside any fixed arc  $|\theta| < v$ .

Here we give the proof to the Theorem 0.2 (check the paper [20] for the related questions).

*Proof. (of the Theorem 0.2).* Since  $\sigma$  belongs to the Steklov class, it belongs to the Szegő class as well and thus the Schur coefficients  $\{\gamma_n\} \in \ell^2$ . In particular,  $\gamma_n \rightarrow 0$  and  $\rho_n \rightarrow 1$ . Take  $z \in \mathbb{T}$ , divide the second equation in (3.2) by  $\phi_n^*$ , and take the absolute value to get

$$\left| \frac{\phi_{n+1}^*}{\phi_n^*} \right| = |\rho_n|^{-1} \cdot \left| 1 - \gamma_n z \frac{\phi_n}{\phi_n^*} \right|$$

Since  $|\phi_n| = |\phi_n^*|$  for  $z \in \mathbb{T}$ , we have

$$\sup_{z \in \mathbb{T}} \left| \left| \frac{\phi_{n+1}(z, \sigma)}{\phi_n(z, \sigma)} \right| - 1 \right| \rightarrow 0, \quad n \rightarrow \infty$$

Iterating, one has

$$\sup_{z \in \mathbb{T}} \left| \left| \frac{\phi_{n+j}(z, \sigma)}{\phi_n(z, \sigma)} \right| - 1 \right| \rightarrow 0, \quad n \rightarrow \infty, \quad j \text{ is fixed} \quad (8.11)$$

Now, suppose (0.10) fails. Then, there is  $\{m_n\} \subseteq \mathbb{N}$  and  $\{z_n\} \in \mathbb{T}$  such that

$$|\phi_{m_n}(z_n)| > C\sqrt{m_n}$$

So, given arbitrary large fixed  $K$ , (8.11) implies

$$|\phi_{m_n+k}(z_n)| > 0.9C\sqrt{m_n}$$

for every  $k : |k| \leq K$  and  $n > n(K)$ . In particular,

$$\sum_{j=0}^{m_n} |\phi_j(z_n)|^2 > (0.9)^2 C^2 K m_n, \quad n > n(K)$$

This, however, contradicts (0.8) as  $K$  is arbitrarily large.  $\square$

## Appendix B.

In this section, we control the phases of various functions we used in the text. Let us start with  $\phi$ , the phase of  $Q_m(e^{i\theta})$ , for  $|\theta| < v$ , where  $v$  is some small, positive, and fixed number.

**Lemma 8.7.** *For any  $\theta \in (-v, v)$ , we have*

$$|\phi'(\theta)| \lesssim m.$$

*Proof.* Recall that (see (3.28))

$$Q_m(z) = \exp \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} C(z, e^{i\xi}) \ln |Q_m(e^{i\xi})| d\xi \right), \quad z \in \mathbb{D} \quad (8.12)$$

and  $\phi(\theta) = \arg Q_m(e^{i\theta})$ , i.e.,

$$\phi(\theta) = \operatorname{Im} \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} C(e^{i\theta}, e^{i\xi}) \ln |Q_m(e^{i\xi})| d\xi \right)$$

where, as before,

$$C(e^{i\theta}, e^{i\xi}) = \frac{e^{i\xi} + e^{i\theta}}{e^{i\xi} - e^{i\theta}}$$

and the integral is taken in principal value. Thus,

$$\phi(\theta) = -\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\cos((\xi - \theta)/2)}{\sin((\xi - \theta)/2)} \ln |Q_m(e^{i\xi})| d\xi$$

This amounts to controlling the Hilbert transform of  $\ln |Q_m(e^{i\xi})|$  since

$$\frac{\cos(\xi/2)}{\sin(\xi/2)} = \frac{2}{\xi} + O(\xi), \quad \left( \frac{\cos(\xi/2)}{\sin(\xi/2)} \right)' = -\frac{2}{\xi^2} + O(1).$$

From the periodicity,

$$\phi(\theta) = -\frac{1}{2\pi} \int_{\theta-\pi}^{\theta+\pi} \frac{\cos((\xi - \theta)/2)}{\sin((\xi - \theta)/2)} \ln |Q_m(e^{i\xi})| d\xi.$$

Changing the variables, we have

$$\phi(\theta) = -\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\cos(\xi/2)}{\sin(\xi/2)} \ln |Q_m(e^{i(\xi+\theta)})| d\xi = -\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\cos(\xi/2)}{\sin(\xi/2)} \cdot \frac{1}{2} (D_m(\xi+\theta) + \ln m) d\xi$$

where

$$D_m(\xi) = \ln \left( \mathcal{G}_m(\xi) + |R_{(m,\alpha/2)}(e^{i\xi})|^2 \right) + \ln m^{-1}.$$

Then,

$$\phi'(x) = -\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\cos(\xi/2)}{\sin(\xi/2)} \frac{1}{2} D'_m(\xi+x) d\xi.$$

We then use the Taylor expansion for

$$\frac{\cos(\xi/2)}{\sin(\xi/2)}$$

and integrate by parts using the periodicity to approximate the integral by the Hilbert transform

$$\begin{aligned} \left| \phi'(x) + \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{D'_m(\xi+x)}{\xi} d\xi \right| &= \left| \int_{-\pi}^{\pi} \left( \frac{\cos(\xi/2)}{\sin(\xi/2)} - \frac{2}{\xi} \right)' \ln \left( \mathcal{G}_m(\xi) + |R_{(m,\alpha/2)}(e^{i\xi})|^2 \right) d\xi \right| \\ &\lesssim \int_{-\pi}^{\pi} \left| \ln \left( \mathcal{G}_m(\xi) + |R_{(m,\alpha/2)}(e^{i\xi})|^2 \right) \right| d\xi \lesssim 1. \end{aligned}$$

The last inequality follows from

$$\ln x < x, \quad x > 1 \quad \implies \quad |\ln x| < 2x + \ln \frac{1}{x}$$

$$\int_{\mathbb{T}} \left( \mathcal{G}_m(\xi) + |R_{(m,\alpha/2)}(e^{i\xi})|^2 \right) d\xi \lesssim 1$$

and

$$\mathcal{G}_m(\xi) + |R_{(m,\alpha/2)}(e^{i\xi})|^2 \gtrsim 1, \quad m > m_0$$

(see Lemma 8.3 for the estimates on  $R_{(m,\alpha/2)}$ ). Therefore, if  $x \in (-v, v)$ , then

$$|\phi'(x)| \lesssim \left| \int_{-\pi}^{\pi} \frac{D'_m(\xi+x)}{\xi} d\xi \right| + 1$$

and

$$|\phi'(x)| \lesssim 1 + m \left| \int_{-m\pi}^{m\pi} \frac{M'_m(t+\hat{x})}{tM_m(t+\hat{x})} dt \right|, \quad M_m(t) = \exp D_m(t/m), \quad \hat{x} = mx$$

$$M_m(t) = \frac{1}{m} \left| Q_m\left(\frac{t}{m}\right) \right|^2$$



For  $M_m$ ,

$$M_m(t) = \frac{\sin^2(t/2)}{m^2 \sin^2(t/(2m))} + \frac{\cos^2(t/2)}{2m^2 \sin^2((t-\pi)/(2m))} + \frac{\cos^2(t/2)}{2m^2 \sin^2((t+\pi)/(2m))} + m^{-1} |R_{(m,\alpha/2)}(e^{it/m})|^2 \quad (8.13)$$

due to (3.25) and (3.27). Thus, we only need to show that

$$I_1(\hat{x}) = \left| \int_{-1}^1 \frac{M'_m(t+\hat{x})}{tM_m(t+\hat{x})} dt \right| \lesssim 1$$

and

$$I_2(\hat{x}) = \left| \int_{1 < |t| < \pi m} \frac{M'_m(t+\hat{x})}{tM_m(t+\hat{x})} dt \right| \lesssim 1$$

uniformly in  $\hat{x} \in [-mv, mv]$ .

Let  $\mathcal{J}_n(\xi)$  denote the sum of the first three terms in (8.13). Then we can rewrite it as follows

$$\begin{aligned} \mathcal{J}_n(\xi) = \sin^2 \frac{\xi}{2} \left( \frac{4}{\xi^2} + \frac{1}{m^2} G\left(\frac{\xi}{2m}\right) \right) + \frac{1}{2} \cos^2 \frac{\xi}{2} \left( \frac{4}{(\xi-\pi)^2} + \frac{1}{m^2} G\left(\frac{\xi-\pi}{2m}\right) \right) + \\ + \frac{1}{2} \cos^2 \frac{\xi}{2} \left( \frac{4}{(\xi+\pi)^2} + \frac{1}{m^2} G\left(\frac{\xi+\pi}{2m}\right) \right) \end{aligned} \quad (8.14)$$

where

$$G(x) = \frac{1}{\sin^2 x} - \frac{1}{x^2}$$

is positive infinitely smooth function defined on  $(-\pi, \pi)$  and  $G(x) \sim 1$  on  $[-a, a] \subset (-\pi, \pi)$ .

Let us start with  $I_2$  and take  $t : |t| < \pi m$ . Therefore, for  $\xi = t + \hat{x}$ , we have  $|\xi| < (\pi + v)m$ .

We will write a lower bound for  $\mathcal{J}_n(\xi)$  for large and for small  $\xi$ .

For large  $\xi$ , i.e.,  $|\xi| \geq c_1 > \pi$ ,  $\frac{|\xi|}{m} \leq c_2 < 2\pi$ , we have:

$$\mathcal{J}_n(\xi) = \frac{4}{\xi^2} \sin^2 \frac{\xi}{2} + \left( \frac{4}{\xi^2} + O(|\xi|^{-3}) \right) \cos^2 \frac{\xi}{2} + O^*\left(\frac{1}{m^2}\right) = \frac{4}{\xi^2} + O(|\xi|^{-3}) + O^*\left(\frac{1}{m^2}\right).$$

For  $\xi \in [-a, a]$  with fixed  $a$ , we get

$$\begin{aligned} \mathcal{J}_n(\xi) &= \frac{4}{\xi^2} \sin^2 \frac{\xi}{2} + \left( \frac{2}{(\xi-\pi)^2} + \frac{2}{(\xi+\pi)^2} \right) \cos^2 \frac{\xi}{2} + O\left(\frac{1}{m^2}\right) \geq \\ &\geq \min\left(\frac{4}{\xi^2}, \frac{2}{(\xi-\pi)^2} + \frac{2}{(\xi+\pi)^2}\right) \left( \sin^2 \frac{\xi}{2} + \cos^2 \frac{\xi}{2} \right) + O\left(\frac{1}{m^2}\right) \gtrsim 1 \end{aligned}$$

Then, for  $|\xi| < (\pi + v)m$ , we have

$$M_m(\xi) \gtrsim \min(1, |\xi|^{-2}) + m^{-1} |R_{(m,\alpha/2)}(e^{i\xi/m})|^2. \quad (8.15)$$

For the derivative of  $M_m(\xi)$ , the representation (8.14) gives an upper bound

$$\begin{aligned} \mathcal{J}'_n(\xi) &= -\frac{8}{\xi^3} \sin^2 \frac{\xi}{2} - \left( \frac{4}{(\xi-\pi)^3} + \frac{4}{(\xi+\pi)^3} \right) \cos^2 \frac{\xi}{2} + O\left(\frac{\|G'\|_\infty}{m^3}\right) + \\ &+ \frac{1}{2} \sin \xi \left[ \frac{4}{\xi^2} - \frac{2}{(\xi-\pi)^2} - \frac{2}{(\xi+\pi)^2} + \frac{1}{m^2} G\left(\frac{\xi}{2m}\right) - \frac{1}{2m^2} G\left(\frac{\xi-\pi}{2m}\right) - \frac{1}{2m^2} G\left(\frac{\xi+\pi}{2m}\right) \right]. \end{aligned}$$

For large  $|\xi|$ , we can write

$$\mathcal{J}'_n(\xi) = -\frac{8}{\xi^3} + O(|\xi|^{-5}) + O\left(\frac{1}{m^3}\right) + O(|\xi|^{-4}) + O\left(\frac{1}{m^4} \|G''\|_\infty\right).$$

For  $\xi \in [-a, a]$  with fixed  $a$ , we again use the smoothness of  $\mathcal{J}_n$ .

$$|\mathcal{J}'_n(\xi)| < \left\| -\frac{8}{\xi^3} \sin^2 \frac{\xi}{2} - \left( \frac{4}{(\xi-\pi)^3} + \frac{4}{(\xi+\pi)^3} \right) \cos^2 \frac{\xi}{2} + \frac{1}{2} \sin \xi \left( \frac{4}{\xi^2} - \frac{2}{(\xi-\pi)^2} - \frac{2}{(\xi+\pi)^2} \right) \right\|_\infty + O\left(\frac{\|G'\|_\infty}{m^3}\right) + O\left(\frac{\pi^2}{8m^4} \|G''\|_\infty\right) \sim 1.$$

Combining these results, we obtain

$$|M'_m(\xi)| \lesssim \frac{1}{(1+|\xi|)^3} + \left| (m^{-1} |R_{(m,\alpha/2)}(e^{i\xi/m})|^2)' \right|, \quad |\xi| \lesssim m. \quad (8.16)$$

First, consider  $\xi : 1 < |\xi| < (\pi + v)m$ . The Lemma 8.3 and (8.15) give

$$M_m(\xi) \gtrsim \frac{1}{\xi^2} + m^{-1} \left| \frac{\xi}{m} \right|^{-\alpha}.$$

Now, it is sufficient to use Lemmas 8.3 and 8.4 to bound the last term in (8.16) as

$$\left| (m^{-1} |R_{(m,\alpha/2)}(e^{i\xi/m})|^2)' \right| \leq \frac{2}{m^2} |B'_m| |B_m| \Big|_{\theta=\frac{\xi}{m}} \lesssim \frac{1}{m^2} \left| \frac{\xi}{m} \right|^{-\alpha/2} \frac{m^{1+\alpha/2}}{|\xi|}.$$

Combining these bounds, we have

$$\left| \frac{M'_m(\xi)}{M_m(\xi)} \right| \lesssim \frac{\frac{1}{|\xi|^3} + \frac{m^{\alpha-1}}{|\xi|^{1+\alpha/2}}}{\frac{1}{\xi^2} + \frac{m^{\alpha-1}}{|\xi|^\alpha}} \leq \frac{\frac{1}{|\xi|^3}}{\frac{1}{\xi^2}} + \frac{\frac{m^{\alpha-1}}{|\xi|^{1+\alpha/2}}}{\frac{m^{\alpha-1}}{|\xi|^\alpha}} \lesssim |\xi|^{-1} + |\xi|^{\alpha/2-1} \quad (8.17)$$

for  $1 < |\xi| < (\pi + v)m$ .

For  $\xi : |\xi| < 1$ , the analogous estimates give

$$\left| \frac{M'_m(\xi)}{M_m(\xi)} \right| \lesssim \frac{1 + m^{\alpha-1}}{1 + m^{\alpha-1}} \lesssim 1. \quad (8.18)$$

Combining (8.17) and (8.18), we get

$$\left| \frac{M'_m(\xi)}{M_m(\xi)} \right| \lesssim (|\xi| + 1)^{-1} + (|\xi| + 1)^{\alpha/2-1} \quad (8.19)$$

which holds uniformly in  $\xi : |\xi| < (\pi + v)m$ . Now, the Cauchy-Schwarz inequality implies the bound for  $I_2$

$$|I_2(\hat{x})| \leq \left( \int_{1 < |t| < \pi m} \frac{dt}{t^2} \right)^{1/2} \left( \int_{|\xi| < (\pi+v)m} \left| \frac{M'_m(\xi)}{M_m(\xi)} \right|^2 d\xi \right)^{1/2} \lesssim 1.$$

Consider  $I_1$ . Apply the Mean Value Formula to rewrite it as

$$|I_1(\hat{x})| = \left| \int_{-1}^1 \frac{1}{t} \left( \frac{M'_n(\hat{x})}{M_n(\hat{x})} + t \left( \frac{M'_n(\xi)}{M_n(\xi)} \right)'_{\xi=\hat{x}_t} \right) dt \right| \lesssim \left\| \frac{M''_m}{M_m} \right\|_\infty + \left\| \frac{M'_m}{M_m} \right\|_\infty^2.$$

The second term was estimated in (8.18) so we only need to control the first one. We use (8.14) and (8.15) to get

$$\left| \frac{M''_m(\xi)}{M_m(\xi)} \right| \lesssim \frac{\frac{1}{(|\xi| + 1)^2} + \left| \left( m^{-1} |R_{(m,\alpha/2)}(e^{i\xi/m})|^2 \right)'' \right|}{\frac{1}{1 + \xi^2} + m^{-1} \left| R_{(m,\alpha/2)}(e^{i\xi/m}) \right|^2}.$$

The estimates from the Lemmas 8.3 and 8.4 in Appendix A can now be used as follows. We have

$$\left| \frac{M_m''(\xi)}{M_m(\xi)} \right| \lesssim 1 + 2 \left| \frac{(R_{(m,\alpha/2)}(e^{i\xi/m}))'}{R_{(m,\alpha/2)}(e^{i\xi/m})} \right|^2 + 2 \left| \frac{(R_{(m,\alpha/2)}(e^{i\xi/m}))''}{R_{(m,\alpha/2)}(e^{i\xi/m})} \right|$$

since  $(R\bar{R})'' = R''\bar{R} + 2R'\bar{R}' + R\bar{R}''$ . For  $\xi : 1 < |\xi| < vm + 1$ , one gets

$$\left| \frac{(R_{(m,\alpha/2)}(e^{i\xi/m}))'}{R_{(m,\alpha/2)}(e^{i\xi/m})} \right| \lesssim |\xi|^{\alpha/2-1}, \quad \left| \frac{(R_{(m,\alpha/2)}(e^{i\xi/m}))''}{R_{(m,\alpha/2)}(e^{i\xi/m})} \right| \lesssim |\xi|^{\alpha/2-1}.$$

For  $\xi : |\xi| < 1$ , we have

$$\left| \frac{(R_{(m,\alpha/2)}(e^{i\xi/m}))'}{R_{(m,\alpha/2)}(e^{i\xi/m})} \right| \lesssim 1, \quad \left| \frac{(R_{(m,\alpha/2)}(e^{i\xi/m}))''}{R_{(m,\alpha/2)}(e^{i\xi/m})} \right| \lesssim 1.$$

This gives a bound

$$\left\| \frac{M_m''}{M_m} \right\|_{\infty} \lesssim 1$$

which ensures

$$|I_1(\hat{x})| \lesssim 1$$

uniformly in  $\hat{x} \in [-mv, mv]$ . The proof is finished.  $\square$

In the next Lemma, we will prove (4.8) and (4.9).

**Lemma 8.8.** *For the functions  $H_n$  and  $\tilde{F}$  introduced in part 2, we have the following bounds*

$$\left| \partial_{\theta} \left( \frac{2 + H_n(1 + \tilde{F})}{2 + \overline{H}_n(1 + \overline{\tilde{F}})} \right) \right| < 0.01n \quad (8.20)$$

and

$$\left| \partial_{\theta} \left( \frac{2 + \overline{H}_n(1 - \tilde{F})}{2 + H_n(1 - \overline{\tilde{F}})} \right) \right| < 0.01n \quad (8.21)$$

provided that  $\rho \ll 1$  and  $n \gg 1$ .

*Proof.* We will only prove (8.20) as the other bound is similar. We have

$$2 + H_n(1 + \tilde{F}) = 2 + (1 - z)(1 - 0.1R_{(m, -(1-\alpha))}(z))(1 + \tilde{C}_n\rho(1 + \varepsilon_n - z)^{-1} + \tilde{C}_n(1 + \varepsilon_n - z)^{-\alpha}).$$

First, notice that  $\tilde{C}_n \rightarrow 1$  as  $\rho \rightarrow 0$  and  $\varepsilon_n \rightarrow 0$  as follows from (4.3). The formula (4.2) for  $H_n$  implies that

$$|H_n(z)| < C(\alpha)|1 - z|, \quad z \in \mathbb{T}$$

Then,

$$\max_{z \in \mathbb{T}} \operatorname{Re} \left( \tilde{C}_n H_n \rho (1 + \varepsilon_n - z)^{-1} \right) \rightarrow 0, \quad \rho \rightarrow 0.$$

Finally,

$$|H_n \tilde{C}_n (1 + \varepsilon_n - z)^{-\alpha}| < \tilde{C}_n |1 + \varepsilon_n - z|^{1-\alpha} (1 + 0.1|1 - z|^{1-\alpha} + o(1))$$

where  $o(1) \rightarrow 0$  as  $m \rightarrow \infty$  because  $R_{(m, -(1-\alpha))}(z) \rightarrow (1 - z)^{1-\alpha}$  uniformly on the circle.

Therefore,

$$|H_n \tilde{C}_n (1 + \varepsilon_n - z)^{-\alpha}| < 2^{1-\alpha} \left( 1 + \frac{2^{1-\alpha}}{10} \right) + o(1) < \sqrt{2} \left( 1 + \frac{\sqrt{2}}{10} \right) + o(1) < 1.8$$

since  $\alpha \in (0.5, 1)$ . Thus,

$$\operatorname{Re}(2 + H_n(1 + \tilde{F})) > 0.2$$

and so

$$|2 + H_n(1 + \tilde{F})| > 0.2$$

for all  $z \in \mathbb{T}$ . Then,

$$\left| \partial_\theta \left( \frac{2 + H_n(1 + \tilde{F})}{2 + \overline{H}_n(1 + \tilde{F})} \right) \right| \leq 2 \left| \frac{\partial_\theta (H_n(1 + \tilde{F}))}{2 + H_n(1 + \tilde{F})} \right| < 10 \left| \partial_\theta (H_n(1 + \tilde{F})) \right|.$$

We have

$$\partial_\theta (H_n(1 + \tilde{F})) = H'_n(1 + \tilde{F}) + H_n \tilde{F}'.$$

The explicit expressions for  $H_n$  and  $\tilde{F}$  give

$$|H_n \tilde{F}'| < C(\alpha) |1 - z| (\rho |1 + \varepsilon_n - z|^{-2} + |1 + \varepsilon_n - z|^{-1-\alpha}) < C(\alpha) (\rho n + n^\alpha), \quad z \in \mathbb{T}$$

and

$$|H'_n(1 + \tilde{F})| < C(\alpha) (1 + |\tilde{F}|) + 0.1 |1 - z| (1 + \tilde{F}) \cdot |R'_{(m, -(1-\alpha))}(z)|.$$

The Lemma 8.6 implies that the third term is bounded by  $Cn^\alpha$  and we have the bound

$$|H'_n(1 + \tilde{F})| < C(\alpha) (1 + \rho n + n^\alpha)$$

uniformly over  $\mathbb{T}$ . Making  $\rho$  small and  $n$  large finishes the proof of (8.20).  $\square$

**Remark.** The estimates in the Lemma above are valid for  $\rho \in (0, \rho_0)$  and  $n > n_0$  where  $\rho_0$  and  $n_0$  both depend on  $\alpha$ .

## REFERENCES

- [1] M. U. Ambroladze, On the possible rate of growth of polynomials that are orthogonal with a continuous positive weight (Russian), *Mat. Sb.* 182 (1991), no. 3, 332–353; English translation in: *Math. USSR-Sb.* 72 (1992), no. 2, 311–331.
- [2] A. I. Aptekarev, V. S. Buyarov, I. S. Dehesa, Asymptotic behavior of  $L^p$ -norms and the entropy for general orthogonal polynomials, *Russian Acad. Sci. Sb. Math.* 1995, 82 (2), 373–395.
- [3] A. I. Aptekarev, J. S. Dehesa, A. Martínez-Finkelshtein, Asymptotics of orthogonal polynomial's entropy. *J. Comput. Appl. Math.* 233 (2010), no. 6, 1355–1365.
- [4] G. E. Andrews, R. Askey, R. Roy, “Special Functions”, Cambridge University Press, 2000.
- [5] B. Beckermann, A. Martínez-Finkelshtein, E. A. Rakhmanov, F. Wielonsky, Asymptotic upper bounds for the entropy of orthogonal polynomials in the Szegő class. *J. Math. Phys.* 45 (2004), no. 11, 4239–4254.
- [6] S. Bernstein, Sur les polynomes orthogonaux relatifs à un segment fini, *Journal de Mathématiques*, (9), 9 (1930), pp. 127–177; 10 (1931), pp. 219–286.
- [7] S. Denisov, On the size of the polynomials orthonormal on the unit circle with respect to a measure which is a sum of the Lebesgue measure and  $p$  point masses, preprint.
- [8] S. Denisov, S. Kupin, On the growth of the polynomial entropy integrals for measures in the Szegő class, *Advances in Mathematics*, Vol. 241, 2013, 18–32.
- [9] P. Duren, *Theory of  $H^p$  spaces*. Dover publications, Mineola, New York, 2000.
- [10] Ya. L. Geronimus, *Polynomials orthogonal on the circle and on the interval*, GIFML, Moscow, 1958 (in Russian); English translation: *International Series of Monographs on Pure and Applied Mathematics*, Vol. 18 Pergamon Press, New York-Oxford-London-Paris, 1960.
- [11] Ya. L. Geronimus, Some estimates of orthogonal polynomials and the problem of Steklov. *Dokl. Akad. Nauk SSSR*, 236 (1977), no. 1, 14–17.
- [12] Ya. L. Geronimus, The relation between the order of growth of orthonormal polynomials and their weight function. *Mat. Sb. (N.S.)* 61 (103), 1963, 65–79.
- [13] Ya. L. Geronimus, On a conjecture of V. A. Steklov. *Dokl. Akad. Nauk SSSR*, 142, 1962, 507–509.
- [14] B. L. Golinskii, The problem of V. A. Steklov in the theory of orthogonal polynomials. *Mat. Zametki*, 15 (1974), 21–32.
- [15] M.E.H. Ismail, *Classical and quantum orthogonal polynomials in one variable*. *Encyclopedia of Mathematics*, 98, Cambridge University Press, 2005.
- [16] A. Kroo, D. Lubinsky, Christoffel functions and universality in the bulk for multivariate orthogonal polynomials. *Canad. J. Math.* 65 (2013), no. 3, 600–620.

- [17] D. Lubinsky, A new approach to universality limits involving orthogonal polynomials. *Ann. of Math. (2)* 170 (2009), no. 2, 915–939.
- [18] A. Mate, P. Nevai, V. Totik, Szegő's extremum problem on the unit circle. *Ann. of Math. (2)* 134 (1991), no. 2, 433–453.
- [19] P. Nevai, Orthogonal polynomials. *Mem. Amer. Math. Soc.* 18 (1979), no. 213.
- [20] P. Nevai, J. Zhang, V. Totik, Orthogonal polynomials: their growth relative to their sums. *J. Approx. Theory* 67 (1991), no. 2, 215–234.
- [21] G. Pólya, G. Szegő, *Problems and theorems in Analysis I*, Berlin-Heidelberg-New York, Springer, 1978.
- [22] E. A. Rahmanov, On Steklov's conjecture in the theory of orthogonal polynomials, *Matem. Sb.*, 1979, 108(150), 581–608; English translation in: *Math. USSR, Sb.*, 1980, 36, 549–575.
- [23] E. A. Rahmanov, Estimates of the growth of orthogonal polynomials whose weight is bounded away from zero, *Matem. Sb.*, 1981, 114(156):2, 269–298; English translation in: *Math. USSR, Sb.*, 1982, 42, 237–263.
- [24] G. Szegő, *Orthogonal Polynomials*, Amer. Math. Soc. Colloq. Publ. **23**, Providence RI, 1975 (fourth edition).
- [25] B. Simon, *Orthogonal polynomials on the unit circle*, volumes 1 and 2, AMS 2005.
- [26] V. A. Steklov, Une methode de la solution du probleme de developement des fonctions en series de polynomes de Tchebysheff independante de la theorie de fermeture, *Izv. Rus. Ac. Sci.*, 1921, 281–302, 303–326.
- [27] P. K. Suetin, V. A. Steklov's problem in the theory of orthogonal polynomials, *Itogi Nauki i Tech. Mat. Anal.*, VINITI, 1977, 15, 5–82 ; English translation in: *Journal of Soviet Mathematics*, 1979, 12(6), 631–682.
- [28] V. Totik, Christoffel functions on curves and domains. *Trans. Amer. Math. Soc.* 362 (2010), no. 4, 2053–2087.
- [29] V. Totik, Asymptotics for Christoffel functions for general measures on the real line. *J. Anal. Math.* 81 (2000), 283–303.
- [30] A. Zygmund, *Trigonometric series*. Third edition, Cambridge University Press, 2002.

UNIVERSITY OF WISCONSIN–MADISON  
 MATHEMATICS DEPARTMENT  
 480 LINCOLN DR., MADISON, WI, 53706, USA  
 DENISOV@MATH.WISC.EDU

KELDYSH INSTITUTE FOR APPLIED MATHEMATICS, RUSSIAN ACADEMY OF SCIENCES  
 MIUSSKAYA PL. 4, 125047 MOSCOW, RUSSIA  
 APTEKAA@KELDYSH.RU